ROYAL AIRCRAFT ESTABLISHMENT

Technical Report 83079

November 1983

THEORY OF DIGITAL IMAGING FROM ORBITAL SYNTHETIC APERTURE RADAR

by

B. C. Barber

Procurement Executive, Ministry of Defence
Farnborough, Hants

UDC 531.7.084.2 : 621.396.965.21

ROYAL AIRCRAFT ESTABLISHMENT

Technical Report 83079

Received for printing 15 November 1983

THEORY OF DIGITAL IMAGING FROM ORBITAL SYNTHETIC APERTURE RADAR

by

B. C. Barber

SUMMARY

Digital synthetic aperture radar (SAR) imaging techniques have previously only been reported in the literature in a fragmentary manner. This article presents a comprehensive review of the theory of digital SAR imaging from Earth orbiting satellites. The digital SAR imaging process is explained, including a discussion of various aspects which are specific to satellite-borne SAR. A number of relevant digital processing techniques are reviewed and it is shown how these techniques may be applied to the processing of digital SAR data. The range migration problem is discussed and various techniques for overcoming it are presented. The paper should be useful not only to the designer of SAR processors, but also to the user of digital SAR data, and images.

Accepted for publication in the International Journal of Remote Sensing.

Departmental Reference: Space 631

LIST OF CONTENTS

1 INTRODUCTION

Synthetic aperture radar imaging of the Earth's surface was proved to be a practical concept by the SEASAT global ocean monitoring satellite launched by NASA in June 1978. Although this satellite remained operational only until October 1978 a considerable quantity of SAR digital data was recorded, and is still being processed at various centres throughout the world.

SEASAT stimulated great interest in orbital SAR and this interest is growing. Since 1978 the Shuttle imaging radar SIR-A has provided more data, although this was optically recorded. The next Shuttle radar mission SIR-B in 1984 will be digitally recorded, and looking further into the future the ESA satellite ERS-1, to be launched in late 1987, will also carry a digital SAR.

Much has been written about the principles of SAR and of optical SAR data processing, see for example the books by Harger[1] and by Hovanessian[2] and the collection of papers edited by Kovaly[3]. Further background information can be found in Refs 13-15. The literature on digital SAR data processing is, however, somewhat fragmented; see for example Refs 4-12. The SAR processor designer needs to have a theory presented as a coherent whole. Others who need such a theory are the image users who must interpret images in the light of such knowledge.

The material presented in this paper is in the nature of an advanced treatment of the theory underlying orbital synthetic aperture radars and provides an introduction to the techniques of digital SAR processing. An attempt has been made to present the theory in a useful form as a coherent whole. In addition, a number of approximations are examined and an attempt made to answer many questions which the author is often asked by workers new to the subject.

An introduction to the basic ideas behind SAR will be found in Refs 1 and 2 and in Refs 13-15. It is assumed in the treatment presented here that the reader is familiar with the contents of these papers.

2 OUTLINE OF THE SAR IMAGING PROCESS

In order to define processes and terms and to fix ideas an outline of the SAR imaging process is presented in this section. In particular, as an illustration, the image of a single point on an absorbing background is considered.

Consider just one pulse, the nth pulse, say, and an echo of unit amplitude from a point on an absorbing background. The received echo is:

$$E(t) = \exp i2\pi\left[f_0(t_n - t_D) + \alpha(t_n - t_D)^2\right] \qquad (1)$$

(see Ref 10).

If the radar receiver outputs a real valued signal then, ignoring a constant phase term

$$E(t) = \cos 2\pi\left[f_0(t_n - t_D) + \alpha(t_n - t_D)^2\right] , \qquad (2)$$

f_0 is the centre frequency of the transmitted range chirp, t_n is the time measured from the centre of the nth pulse, t_D is the round trip time from radar to point to radar for this particular pulse. If $a(t)$ is the range of the point from the radar then $t_D = 2a(t)/c$ where c is the velocity of light. In this section $a(t)$ is assumed constant for one pulse but varies from pulse to pulse. This is the 'stop-start' approximation whereby the radar platform is modelled as being stationary while a pulse is being transmitted and received; t is the overall time coordinate.

In equations (1) and (2) the transmitted chirp has been put implicitly into time symmetric form:

$$A(t) = \exp i2\pi t_n(f_0 + \alpha t_n) \qquad -T/2 \leqslant t_n \leqslant T/2 \qquad (3)$$

where T is the pulse duration and α = bandwidth/2T .

A number of interlinked approximations are implicitly made in this section. Briefly these are:

(1) The 'stop-start' approximation already described.

(2) When a signal is received as a result of scattering from a target and there is a relative velocity between transmitter/receiver and the target the frequency of the received signal will be changed by a scaling factor β (the Doppler effect). The scaling factor β is taken to be $\beta = 1 + 2v/c$ where v is the relative velocity of the radar with respect to the target. The signal spectrum is transformed $s(f) \rightarrow s(f/\beta)/\sqrt{|\beta|}$ and f/β is taken to be $f(1 - 2v/c)$ and $\sqrt{|\beta|} = 1$. This is justified by the fact that $v \ll c$.

(3) Although for a spacecraft SAR the signals received by the radar were transmitted a few milliseconds earlier when the radar was at another position it is assumed that the round trip range and Doppler frequency shift are those corresponding to the time of reception. An examination of this approximation for the SEASAT SAR shows that the main effects are an error of order 1 m in the range direction resulting in a totally negligible defocussing in the azimuth direction and an image shift of order 10 m in the azimuth direction with no corresponding defocussing.

These approximations are associated with the fact that during the time interval for which a point on the ground is illuminated by a pulse (33.9 μs for

SEASAT) the phase change caused by the varying range is very small. It can then be shown that each pulse samples the phase of the point, that is $4\pi a(t)/\lambda$ for the nth pulse. This whole question is examined in section 3 where it is shown that the 'stop-start' approximation is entirely acceptable.

2.1 Range compression

On receipt of each pulse the radar receiver coherently mixes the rf centre frequency down to the 'offset video frequency' f_1. The offset frequency is usually chosen to be one quarter of the sampling frequency of the analogue-to-digital converter. The reason for this will become clear in section 6.1. For SEASAT this frequency was approximately 11.38 MHz. The received echo is then:

$$E(t) = \cos 2\pi\left[f_1 t_n - f_0 t_D + \alpha(t_n - t_D)^2\right] \quad . \tag{4}$$

It should be made clear at this stage that it has been assumed here that the signal received and input to the subsequent processing is real-valued. It is entirely possible for a SAR receiver to produce a complex signal with a single sideband via a quadrature filter (see Ref 16, p 119). In that case an offset video frequency of zero would be chosen. However, this is not usual since it is easier to sample the real valued signal and convert to a complex valued signal with a single sided spectrum in the subsequent digital processing - see section 5.1.

Range compression is performed on each pulse by correlating the pulse against a replica of the transmitted pulse translated to the offset video band. The peak of the correlation occurs at the round trip delay time for this particular pulse, *ie* at $\tau_R = t_D$ where τ_R is the correlation variable. In the correlation which follows the time symmetric forms of transmitted and received pulses are used and the start of transmission occurs at $t_n = -T/2$ so that $t_n = 0$ coincides with the centre of the transmitted pulse. Correlating (4) against (3) over all values of t_n from $t_D - T/2$ to $t_D + T/2$ gives for the range point spread function

$$\begin{aligned} g_R = {} & \tfrac{1}{2} \int_{t_D-T/2}^{t_D+T/2} \exp i2\pi\left[f_1 t_n - f_0 t_D + \alpha(t_n - t_D)^2\right] \times \\ & \qquad \times \exp - i2\pi\left[f_1(t_n - \tau_R) + \alpha(t_n - \tau_R)^2\right] dt_n \\ & + \tfrac{1}{2} \int_{t_D-T/2}^{t_D+T/2} \exp - i2\pi\left[f_1 t_n - f_0 t_D + \alpha(t_n - t_D)^2\right] \times \\ & \qquad \times \exp i2\pi\left[f_1(t_n - \tau_R) + \alpha(t_n - \tau_R)^2\right] dt_n \quad , \end{aligned} \tag{5}$$

and so

$$g_R = \cos 2\pi\left[- f_0 t_D + f_1 \tau_R - \alpha(t_D - \tau_R)^2\right] \frac{\sin \pi\Delta F_R(t_D - \tau_R)}{\pi\Delta F_R(t_D - \tau_R)} \qquad (6)$$

where ΔF_R = the chirp (video) bandwidth = $2\alpha T$.

Equation (6) assumes continuous functions. In reality these functions are sampled, and the correlation is performed in the frequency domain via an FFT. This is explained in section 5.4. The result of the correlation (6) is, of course, the same if the correlation is performed in the frequency domain via Fourier transforms, although the range point spread function (6) is slightly different when performed on sampled data and is periodically repeated. At this stage of the processing the cosine term in equation (6) is turned into an exponential. The reason for this is explained in section 5.1. If (6) is Fourier transformed with respect to τ_R one obtains an approximately rectangular spectrum (it is a Fresnal integral) of width ΔF_R centred on f_1 multiplied by $\exp i2\pi t_D(f_1 - f_0)$ and an approximately rectangular spectrum of width ΔF_R centred on $-f_1$ multiplied by $\exp - i2\pi t_D(f_1 - f_0)$. As was mentioned previously the range correlation is carried out in the frequency domain and the simple step of setting all the negative frequency samples to zero gives only the positive half of the spectrum. If this is inverse Fourier transformed it will be seen that instead of (6) we have:

$$g_R = \exp i2\pi\left[- f_0 t_D + f_1 \tau_R - \alpha(t_D - \tau_R)^2\right] \frac{\sin \pi\Delta F_R(t_D - \tau_R)}{\pi\Delta F_R(t_D - \tau_R)} \quad . \qquad (7)$$

At the same time the offset video frequency f_1 is mixed to zero by the simple procedure of moving all the frequency samples (see section 8). The resulting range point spread function is then:

$$g_R = \exp - i2\pi\left[f_0 t_D + \alpha(t_D - \tau_R)^2\right] \frac{\sin \pi\Delta F_R(t_D - \tau_R)}{\pi\Delta F_R(t_D - \tau_R)} \quad . \qquad (8)$$

In passing it should be pointed out that these expressions for g_R have been multiplied by a constant ($= A^{-\frac{1}{2}}$, where A is the time bandwidth product) in order to normalize them.

2.2 Azimuth compression

The next main step in the processing is azimuth correlation (although in practice, between range correlation and azimuth correlation there are the very

important steps of corner turning and range migration correction. These are computational problems and are examined in sections 2.3 and 10). Azimuth correlation is essentially the same process as range correlation but in this case the frequency modulation results from the Doppler effect and the change in t_D from pulse to pulse. The change in round trip time t_D gives a rotating vector via the phase term in equation (8). The rate at which it rotates represents the Doppler frequency. In a sense, then, each pulse samples the phase of the point and over a number of pulses the rate of change of phase is built up, and one can then meaningfully refer to a frequency. The change in phase from pulse to pulse is a linear function of t_D in equation (8) because as t_D varies from pulse to pulse the value of τ_R for which the samples are selected from each pulse for azimuth compression tracks t_D with an arbitrary offset. The quadratic function of t_D in (8) therefore remains constant. This is a consequence of range migration correction and is explained in more detail in section 2.3. The remaining phase term in (8) gives the phase history of the point. Now we may write

$$t_D = \frac{2a(t)}{c} = \frac{2}{c}\left(a_0 + a_1 t + a_2 t^2 + \ldots\right) \tag{9}$$

where the range to the target point has been expanded in a Taylor series about $t = 0$ (*ie* the centre of the synthetic aperture). That is, at $t = 0$ we have

$$a = a_0 + a_1 t + a_2 t^2 + \ldots \tag{10}$$

where a_0 is the slant range from the radar to the point, a_1 is the slant range velocity and a_2 is one half of the slant range acceleration. Usually only terms up to the quadratic term are taken into account but sometimes, for extra precision, the cubic term is used as in the RAE processor. The range polynomial (10) is calculated for the general case of an elliptic orbit and rotating Earth in section 4 where it is shown that the finite value of a_1 at the centre of the aperture is a result of radar squint, orbit eccentricity and Earth rotation.

We now correlate in the azimuth direction and at a temporal slant range displacement from t_D of τ_R. Note that τ_R is *not constant*. This is the so called range migration effect. Firstly, because the Earth rotates, a point on it is taken through many range samples (range gates) possibly several dozen for a high resolution spacecraft SAR. Secondly, for a high resolution spacecraft SAR the geometry is such that an arc of constant slant range cannot usually be considered 'straight' and the curvature can cause the point to move through a number of range gates. These two effects are known as range walk and range curvature. Together they constitute range migration and cause serious problems for the SAR processor designer. These effects are considered further in sections 2.3 and 13.

In passing, note that the range time coordinate τ_R is not constant, it tracks t_D and has an arbitrary constant offset from it corresponding to the displacement of the line of data being used for azimuth compression from the centre of the range point spread function $|t_D - \tau_R|$.

Some approximations are now made in order to make the azimuth correlation process clearer although it should be noted that these are not actually made in a working digital processor. For a small change in slant range the first order changes to the range polynomial are Δa_0 in a_0 , $\Delta a_0(a_1/a_0)$ in a_1 and $\Delta a_0(a_2/a_0)$ in a_2 . For lengths comparable with the point spread function scale it can easily be shown that the changes to a_1 and a_2 are negligible, so that only the change Δa_0 in a_0 , need be considered. Thus expanding a about $a = a_0$ and $t = 0$ the following terms are retained:

$$a'(t) = a_0 + \Delta a_0 + a_1 t + a_2 t^2 \tag{11}$$

where

$$\Delta a_0 = \tau_R \frac{c}{2} \ . \tag{12}$$

In addition, for a displacement d in the direction of satellite motion, there will be a shift in the azimuth time coordinate τ_A equal to d/v_p where v_p is the local velocity of the satellite.

There is ample scope for confusion here. Conventionally, the azimuth direction is defined as the direction of the satellite motion. This has the advantage of being constant over the whole swath width. Strictly, however, it ought to be defined as the direction of the resultant of v_p and $da(t)/dt$ at the relevant point. This latter definition is used here and the two directions are called the 'along track direction' and the 'azimuth direction'. The true azimuth direction, then, varies over the swath and if an image of a bright point is observed it will be seen that the sidelobes in the azimuth direction are aligned along a slight curve which does not intersect the range direction at right angles. This is illustrated in Fig 1.

The actual phase history of the point is $\exp(-i2\pi f_0 t_D)$ with t_D given by (9) and the phase history against which it is correlated by the processor is $\exp(-i2\pi f_0 t'_D)$ with t'_D given by:

$$t'_D = \frac{2}{c}\left[a_0 + \Delta a_0 + a_1(t + \tau_A) + a_2(t + \tau_A)^2\right] \ . \tag{13}$$

It should be noted that these various approximations have been introduced here in order to make this explanation clear. In practice digital processors compute the phase history directly from the spacecraft orbit, the Earth rotation and the geometry. Correlating (8) against $\exp(-i2\pi f_0 t_D')$, then, gives:

$$g(\tau_R,\tau_A) = \exp - i2\pi\alpha(t_D - \tau_R)^2 \frac{\sin \pi\Delta F_R(t_D - \tau_R)}{\pi\Delta F_R(t_D - \tau_R)} \times$$

$$\times \int_{-T_A/2}^{T_A/2} \exp - i \frac{2\pi}{c} 2f_0\left[a_0 + a_1 t + a_2 t^2\right] \times$$

$$\times \exp i2\pi \frac{2f_0}{c}\left[a_0 + \Delta a_0 + a_1(t - \tau_A) + a_2(t - \tau_A)^2\right]dt \qquad (14)$$

where a synthetic aperture interval of $\pm T_A/2$ centred on zero has been taken. An asymmetric aperture simply introduces an additional phase term. So

$$g(\tau_R,\tau_A) = \exp - i2\pi\alpha(t_D - \tau_R)^2 \exp i2\pi \frac{2f_0}{c}\left(\Delta a_0 + a_1\tau_A + a_2\tau_A^2\right) \times$$

$$\times \frac{\sin \pi\Delta F_R(t_D - \tau_R)}{\pi\Delta F_R(t_D - \tau_R)} \frac{\sin \pi\Delta F_A\tau_A}{\pi\Delta F_A\tau_A} \qquad (15)$$

ΔF_A is the Doppler bandwidth. The point spread function $g(\tau_R,\tau_A)$ can be expressed in lengths in azimuth (x) and slant range (y) by means of $\tau_A = x/v_p$ and $\tau_R = 2y/c$. The point spread function in (15) has been scaled by multiplying by $A^{-\frac{1}{2}}B^{-\frac{1}{2}}$ where A and B are the range and azimuth time bandwidth products.

2.3 Range migration

It will be observed that the position of the range point spread function g_R (see equation (8)) is a function of time because t_D is a function of time as it must be to obtain a Doppler frequency shift on which the synthetic aperture principle depends. If this shift is much smaller than the range sample separation then its effect on the azimuth correlation integral can be ignored: for example most aircraft synthetic aperture radar systems are designed so that this is the case. However, as pointed out in the last section, it is not possible to do this in the case of an orbital SAR for a high resolution imaging system. The effect of the temporal dependence of the range point spread function is to couple the range and

along track directions in the azimuth correlation integral. If the azimuth correlation integral is performed along the 'true' azimuth direction then the range and azimuth correlations are uncoupled. The 'true' azimuth direction was defined in the last section in one way, and in this section it is implicitly defined in another, but closely related and equivalent way. Returning to the azimuth correlation (14) and putting the time dependent range point spread function inside the integral we obtain

$$g(\tau_R,\tau_A) = \int_{-T_A/2}^{T_A/2} \exp - i2\pi\alpha(t_D - \tau_R) \frac{\sin \pi\Delta F_R(t_D - \tau_R)}{\pi\Delta F_R(t_D - \tau_R)} \times$$

$$\times \exp - i \frac{2\pi}{c} 2f_0\left[a_0 + a_1 t + a_2 t^2\right] \times$$

$$\times \exp i2\pi \frac{2f_0}{c} \left[a_0 + \Delta a_0 + a_1(t - \tau_A) + a_2(t - \tau_A)^2\right]dt \ . \quad (16)$$

The range and along track components are uncoupled by performing an additional convolution operation along the line defined by

$$t_D = \frac{2}{c}\left(a_0 + a_1 t + a_2 t^2 \ \ldots\right)$$

in range sample time and azimuth time space. Thus if $t_R = t_D - \tau_R$:

$$g(t_R,\tau_A) = \int_{-T_A/2}^{T_A/2} \int_{t_D - T/2}^{T_D + T/2} \delta(t_D + t_R - \tau_R)d\tau_R \times$$

$$\times \exp - i2\pi\alpha(t_D - \tau_R) \frac{\sin \pi\Delta F_R(t_D - \tau_R)}{\pi\Delta F_R(t_D - \tau_R)} \times$$

$$\times \exp - i \frac{2\pi}{c} 2f_0\left[a_0 + a_1 t + a_2 t^2\right] \times$$

$$\times \exp i2\pi \frac{2f_0}{c} \left[a_0 + \Delta a_0 + a_1(t - \tau_A) + a_2(t - \tau_A)^2\right]dt \quad (17)$$

where the range compressed data spans the time interval $t_D \pm T/2$. It will be noted that the additional operation must be carried out *before* azimuth correlation.

Usually the convolution is carried out by 'data selection' using more or less approximate methods; this is explained further in section 14.

3 THE STOP-START APPROXIMATION

In section 2 the range compression process and subsequent azimuth compression were based on the 'stop-start' approximation in which the delay time t_D was taken to be constant over a pulse width. This approximation, and others closely linked with it, outlined in section 2, are now examined in more detail in this section. In digital processing of any kind and in particular in digital SAR processing it is necessary to have a very clear analysis of all the details of the imaging process including the approximations since a mathematical model must be constructed. Failure to do this can lead to unsatisfactory results and unexpected problems.

3.1 The delay time

Let the radar be at rest in a reference frame S with coordinates (x,y,z,t) and let the target point be at rest in a frame S' with coordinates (x',y',z',t'). There is a relative velocity between the frames and the two frames are related via a Lorentz transformation (see Ref 17, p 110 and also p 132). For the moment any accelerations are ignored. A radar operates by sending out a pulse of electromagnetic waves at time t_1 say, and receiving the pulse at t_3 both times being measured in frame S at rest relative to the radar. The time at which the pulse is estimated to arrive at the target point by an observer at the radar is then $(t_3 + t_1)/2$; hence

$$t_2 = (t_1 + t_3)/2 \quad . \tag{18}$$

A different time is measured by the observer at the target since to him it appears that the clocks in the radar reference frame are slow. In the Lorentz transformation which connects an event in each frame the relative velocity appears as v^2/c^2 . An Earth-orbiting spacecraft may have a velocity of several kilometres per second, say 6.7 km/s at a (typical) orbital height of 800 km and $v^2/c^2 \simeq 5 \times 10^{-10}$. This is generally not significant for a low frequency SAR such as the L-band SEASAT but note that at a wavelength of 3 cm the carrier frequency is 10^{10} Hz and hence the second order term in v/c leads to a Doppler shift of 5 Hz or so which might be measurable. Thus with future X-band or higher frequency orbiting SARs the effect may have to be taken into account as will be explained later in the section.

In the analysis which follows v^2/c^2 is very small and so the factor $\sqrt{1 - v^2/c^2}$ in the Lorentz transformation from one frame to another will always be taken as unity. Then the analysis will be correct up to first

order terms in v/c . Within this approximation, time can be taken as absolute and the analysis can be performed in 'absolute' space.

Let the position vector of the radar be $\bar{r}_1(t)$ and that of the target point $\bar{r}_2(t)$ in any suitable coordinate system. An electromagnetic wave leaves the radar at time t_1 and travels towards the target arriving there at time t_2, is reflected and arrives back at the radar at t_3 . The distance travelled by the wave from radar to target

$$d_1 = |\bar{r}_1(t) - \bar{r}_2(t)| \tag{19}$$

and from target to radar

$$d_2 = |\bar{r}_2(t) - \bar{r}_1(t)| \quad . \tag{20}$$

Hence the total time of the wavefront from radar to target to radar is

$$t_3 - t_1 = \frac{1}{c}\left[|\bar{r}_2(t_2) - \bar{r}_1(t_3)| + |\bar{r}_1(t_1) - \bar{r}_2(t_2)|\right] \quad . \tag{21}$$

Since we are dealing with differences it is easier to redefine the coordinate system so that the origin is at the target point, *ie* $\bar{r}_2(t) \equiv 0$. Then

$$\bar{r}_1(t) \equiv \bar{r}(t)$$

and so

$$t_3 - t_1 = \frac{1}{c}\left[r(t_2) + r(t_1)\right] \quad . \tag{22}$$

In equation (22) the time of reception corresponding to t_3 is t_n and the delay time is t_D . Hence

$$t_D = \frac{1}{c}\left[r(t_n) + r(t_n - t_D)\right] \tag{23}$$

t_D must now be obtained explicitly by solving this equation. Fortunately t_D is usually small (around 6 ms for a slant range of 850 km) and $\bar{r}(t)$ is a smooth function so that $r(t_n - t_D)$ can be expanded as a Taylor series around $r(t_n)$. Now,

$$r(t_n - t_D) = \left|\bar{r}(t_n) - \bar{v}(t_n)t_D + \bar{a}(t_n)t_D^2/2 + \ldots\right| \tag{24}$$

where $\bar{v} = \partial\bar{r}/\partial t$ and $\bar{a} = \partial\bar{v}/\partial t$, which gives

$$r(t_n - t_D) = r(t_n) - \hat{r}(t_n) \cdot \bar{v}(t_n) t_D + O\left(t_D^2\right) \tag{25}$$

$\hat{r}$ being the unit vector in the direction of $\bar{r}$. It is not necessary to calculate the second order term since the eventual expression for t_D will only be accurate up to first order in v/c . So we have from equations (23) and (25)

$$t_D = \frac{1}{c}\left[2r(t_n) - \hat{r}(t_n) \cdot \bar{v}(t_n) t_D + O\left(t_D^2\right)\right] \tag{26}$$

which gives

$$t_D = \frac{2r(t_n)}{c}\left[1 - \frac{\hat{r}(t_n) \cdot \bar{v}(t_n)}{2c} + O\left(\frac{v^2}{c^2}\right)\right] . \tag{27}$$

A more elaborate analysis including the second order terms confirms this equation and shows that the acceleration terms are of order v^2/c^2 and thus negligible. The angle between $\hat{r}(t)$ and $\bar{v}(t)$ is the azimuth angle ϕ , equal to $\pi/2$ when the radar is broadside on to the target point. One might expect the term $\hat{r}(t_n) \cdot \bar{v}(t_n)/c$ to be negligibly small and this turns out to be the case. We have

$$t_D = \frac{2r(t_n)}{c}\left[1 - \frac{v(t_n)}{2c}\cos\phi(t_n) + O\left(\frac{v^2}{c^2}\right)\right] . \tag{28}$$

Note that at this stage the time delay t_D is still a function of time t_n via $r(t_n)$, $v(t_n)$ and $\phi(t_n)$.

Define τ_0 to be the time at the centre of the received pulse when t_D is assumed constant, *ie* if $v = 0$ and corresponds to the fixed t_D of the stop-start approximation. The delay time at τ_0 is then

$$t_D(\tau_0) = \frac{2r(\tau_0)}{c}\left[1 - \frac{v(\tau_0)}{2c}\cos\phi(\tau_0) + O\left(\frac{v^2}{c^2}\right)\right] . \tag{29}$$

We require the time delay at time offsets from τ_0 up to $\pm T/2$. The time delay at time τ is then

$$t_D(\tau) = t_D(\tau_0) + t_D(\tau_0)\tau + t_D(\tau_0)\tau^2/2 + \ldots . \tag{30}$$

Differentiating (28) to give the coefficients in this Taylor expansion about τ_0 gives terms involving $2r(t_n)/c$ and its derivatives and $vr(t_n)\cos\phi(t_n)/c^2$

and its derivatives. The maximum rate of change of ϕ occurs for the broadside mode and is typically a few tens of mrad/s maximum. Also, as always, $v \ll c$ and the leading terms only are taken in the expansion to give

$$\dot{t}_D(\tau_0) = \frac{2\dot{r}(\tau_0)}{c} \quad \text{and} \quad \ddot{t}_D(\tau_0) = \frac{2\ddot{r}(\tau_0)}{c}$$

and hence

$$t_D(\tau) = \frac{2}{c}\left[r(\tau_0) + \dot{r}(\tau_0)\tau + \ddot{r}(\tau_0)\frac{\tau^2}{2}\right] . \tag{31}$$

3.2 Range compression with varying time delay

Let

$$t_D(\tau) = \tau_0 + \tau_1\tau + \tau_{\varepsilon 1} \tag{32}$$

where

$$\tau_{\varepsilon 1} = \frac{\tau_2\tau^2}{2!} + 0(\tau^3) \tag{33}$$

and

$$t_D^2(\tau) = \tau_0^2 + 2\tau_0\tau_1\tau + \tau_{\varepsilon 2} \tag{34}$$

where

$$\tau_{\varepsilon 2} = \frac{2\tau_0\tau_2\tau^2}{2!} + \tau_1^2\tau^2 + 0(\tau^3) , \tag{35}$$

and from (31)

$$\tau_0 = \frac{2r(\tau_0)}{c} , \quad \tau_1 = \frac{2\dot{r}(\tau_0)}{c} , \quad \tau_2 = \frac{2\ddot{r}(\tau_0)}{c} .$$

In addition

$$g_R = \int_{-T/2}^{T/2} \exp - i2\pi\left[f_0 t_D - f_1\tau_R + \alpha(t_D - \tau_R)^2 + 2\alpha\tau(t_D - \tau_R)\right]d\tau . \tag{36}$$

This is obtained by rearranging the first integral in (5), the second is ignored for the purposes of the present analysis. Substituting from (32) and (34) we obtain

$$g_R = \exp - i2\pi\left[f_0\tau_0 - f_1\tau_R + \alpha(\tau_R - \tau_0)^2\right] \times$$

$$\times \int_{-T/2}^{T/2} \exp - i2\pi\left[f_0\tau_{\varepsilon 1} - \alpha\tau_{\varepsilon 2} + 2\alpha\tau_1\tau^2 + 2\alpha\tau\tau_{\varepsilon 1} + 2\alpha\tau_0\tau_{\varepsilon 1}\right] \times$$

$$\times \exp - i2\pi\ 2\alpha\tau\left[(\tau_0 - \tau_R) + \frac{f_0\tau_1}{2\alpha}\right] d\tau \quad . \tag{37}$$

The first exponential under the integral contains quadratic and higher powers of τ and the second exponential contains the linear terms in τ. The non-linear powers of τ result in defocussing and the linear terms simply result in a shift in the position of the point spread function in the range direction. The quadratic phase term is slowly varying since the non-linear terms are small. The linear phase term is also slowly varying in the neighbourhood of the peak of the point spread function since $\tau_0 \simeq \tau_R$ near the peak. The non-linear term is

$$f_0\tau_2\frac{\tau^2}{2} + \alpha\tau_1\tau^2(2 - \tau_1) + O(\tau^3) \quad . \tag{38}$$

Note that $\tau_1 = 2\dot{r}(\tau_0)/c \ll 2$ since $\dot{r} \leqslant v$ and $v \ll c$. Also $2\alpha\tau_1 \gg f_0\tau_2/2$ since α is typically of the order of 10^{12} Hz/s. The largest non-linear term is then $2\alpha\tau_1\tau^2$. In section 4 it is demonstrated that a typical value of $\dot{r}$ for a broadside looking orbital SAR at an altitude of 800 km is of order 100 m/s. The quadratic phase change over the correlation is then of order 1 mrad for a pulse width of the order of tens of microseconds. This magnitude of phase change is completely negligible.

The first order effect of the stop-start approximation is a shift in the position of the point spread function by $f_0\tau_1/2\alpha$, which, for SEASAT, is about 1.5 ns or 0.45 m in slant range. This error, however, could be much greater for a radar operating at a higher frequency. Also, in a SAR operating in a squint mode τ_1 could easily be an order of magnitude greater. A shift in the position of the point spread function, however, causes no real problems unless the shift varies over the synthetic aperture by more than, say one half resolution cell. τ_1 does vary in the azimuth direction but not by enough to cause problems with the approximation considered here. Thus in conclusion the stop-start approximation is valid to a high degree of precision for orbital synthetic aperture radars.

4 THE RANGE POLYNOMIAL

In section 2.2, equation (10), the range from radar to point was expanded as a function of time:

$$a(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 \ldots \quad . \tag{10}$$

In this section the terms in the expansion are calculated for a general elliptic orbit, and various approximations considered.

It seems appropriate to mention here that there are at least three modes in which a SAR may operate. These are the normal mode, the squint mode and the spotlight mode. In the normal mode the radar looks out sideways to the track and perpendicular to it. The squint mode is similar except that the radar is permanently squinted either forwards or backwards to the normal mode. In the spotlight mode the radar squint angle is varied continuously as the radar passes a target so that the target is constantly illuminated, and hence the synthetic aperture length is no longer limited by the footprint width. A very large aperture can then, in principle, be synthesised. The squint angle is varied by varying the pitch angle and/or the yaw angle of the spacecraft depending on the angle of incidence of the radar.

The analysis which follows is applicable (within limits) to all three modes. Before calculating the range polynomial, however, some considerations concerning the choice of orbit for SAR spacecraft are presented.

4.1 Orbits for SAR satellites

Circular or nearly circular orbits are arguably the best choice for radar remote sensing satellites. There are several reasons for this. First, if an approximately constant height is maintained above the Earth's surface then the Doppler bandwidth will remain constant and this leads to a constant pulse recurrence frequency. The prf is one of the main design parameters for a synthetic aperture radar and if it can be kept constant this greatly simplifies the design of the radar system. Further discussion of this point belongs to the design of the radar rather than the processing and is not pursued here.

Secondly, if the radar data are required to be processed into images soon after the satellite pass then nominal (and possibly inaccurate) orbital elements will have to be used in the processing. Refined orbital parameters are not available until several days have elapsed because of the necessity of measuring the orbit at several ground stations around the Earth. The inaccuracy in the processing due to the use of nominal orbital elements seems to be minimised if a circular orbit is employed, or if the orbit is frozen in some way.

It could be argued that it might be desirable to have a slightly elliptical orbit so as to maintain (to first order) a constant height above the Earth. The orbit apogee would then lie on the equator, the Earth being a flattened spheroid bulging at the equator. This argument was advanced in Ref 19 in connection with the SEASAT altimeter and such an orbit has an advantage because if the argument of perigee is 90° (which, of course, it is if the apogee lies on the equator) then, in principle, it is possible to choose an inclination and eccentricity such that the orbit is frozen. That is to say such that the perigee does not precess. Normally, the perigee processes in the orbit plane by a few degrees per day due to the Earth's oblateness. The reader is referred to Ref 19 for a fuller account.

Another effect of the Earth's oblateness is to precess the plane of the orbit by a few degrees per day. The rate of precession $\dot{\Omega}$ is given approximately by:

$$\dot{\Omega} = -9.97\left(\frac{R}{a}\right)^{7/2} \frac{\cos i}{(1-e^2)^2} \quad \text{degrees/day} \tag{39}$$

see Ref 18, equation 1.1. It will be observed that this is a function of semi-major axis a, eccentricity e and angle of inclination i.

This effect has some importance to spacecraft designs because the rate of precession can be chosen to be such that the solar panels which power most satellites continuously face the Sun. This is a Sun-synchronous orbit and has a rate of precession of approximately 360° in 365 days. Not too much weight should be placed on this requirement, however, because quite a large variation on a perfect Sun-synchronous orbit is acceptable and not all spacecraft are powered by solar cells. If one were designing for a synchronous orbit then once the height of the satellite above the ground is chosen (this is a function of radar maximum power, area of coverage, maximum data rate capacity etc) then the inclination angle i follows for a given eccentricity. The required inclination is greater than 90° since $\dot{\Omega}$ must be positive and such orbits are known as retrograde orbits. For example in the case of SEASAT i was slightly greater than 108°. Of course, if the radar platform is not dependent on the Sun for its power (as is the case for the Shuttle for example) then the orbit is not constrained in this manner.

The above discussion is a somewhat simplified account of the choice of orbit for a SAR satellite. The actual details are complicated and further details of the choice of orbit for SEASAT will be found in Ref 19.

In the analysis which follows a typical orbit is assumed to be as in Fig 2. The orbit is shown projected on to a spherical surface which is 'fixed' relative to the background stars. The direction labelled Υ is the vernal equinox (the intersection of the Earth's equator and the plane of the ecliptic) and the centre of the surface coincides with the centre of the Earth. The point P is the point on the Earth's surface for which the range polynomial is to be computed. The latitude of the point is measured in geocentric coordinates and is constant, but the longitude is constantly increasing due to the Earth's rotation. The point A is the projection of the real antenna focus on to the celestial sphere.

4.2 Latitude and longitude changes of radar and target

Let the point P in Fig 2 have a geocentric latitude and longitude measured on the fixed surface of ψ and ϕ. Let R be the distance of the focus of the (real) radar antenna from the Earth's centre and r the distance of P from the Earth's centre. Then if a is the distance between the radar antenna focus and the point P

$$a^2 = R^2 + r^2 - 2Rr \cos\theta \qquad (40)$$

where θ is the angle $A\hat{O}P = A\hat{O}D$ in Fig 2. θ is a function of time due to the motion of the satellite and the rotation of the Earth as indicated in Fig 2. R is also a function of time due to the non-circular orbit, but r is constant.

In Fig 2 D is the projection of r on to the celestial sphere upon which the orbit has been projected, and the arc AD passing through the projection of the antenna focus A and D has on a great circle centre O. The meridians passing through A and D are also great circles and triangle ADN, N being the north pole is a spherical triangle. Let the latitude and longitude of A be Ψ and Φ in geocentric coordinates. Then from ΔADN we have:

$$\cos\theta = \sin\Psi \sin\psi + \cos\Psi \cos\psi \cos(\Phi - \phi) \qquad (41)$$

which gives θ in terms of the latitudes and longitudes of $A(\Psi,\Phi)$ and $P(\psi,\phi)$.

The purpose of the analysis here is to compute the coefficients $a_0 \ldots a_3$ in the range polynomial

$$a = a_0 + a_1 t + a_2 t^2 + a_3 t^3 \ldots$$

in the neighbourhood of $t = 0$. This involves differentiating (40) successively and setting $t = 0$. Hence explicitly we have:

$$a_0 = a \quad \text{at} \quad t = 0 \tag{42}$$

$$a_0 a_1 = \dot{R}_0(R_0 - r\cos\theta_0) - R_0 r \frac{d}{dt}(\cos\theta_0) \tag{43}$$

$$2a_0a_2 + a_1^2 = \ddot{R}_0(R_0 - r\cos\theta_0) + \dot{R}_0\left[\dot{R}_0 - 2r\frac{d}{dt}(\cos\theta_0)\right] - R_0 r \frac{d^2}{dt^2}(\cos\theta_0) \tag{44}$$

$$6a_0a_3 + 6a_1a_2 = \dddot{R}_0(R_0 - r\cos\theta_0) + \ddot{R}_0\left[2\dot{R}_0 - 3r\frac{d}{dt}(\cos\theta_0)\right] + \dot{R}_0\left[\ddot{R}_0 - 3r\frac{d^2}{dt^2}(\cos\theta_0)\right] - R_0 r \frac{d^3}{dt^3}(\cos\theta_0) . \tag{45}$$

The subscript (0) indicates that the variable is taken at $t = 0$. When the orbit is circular $\dot{R}_0$ and higher derivatives are zero and the equations simplify greatly. To proceed further it is necessary to compute derivatives of $\cos\theta$ from (41) in the region of $t = 0$.

Let Fig 2 show the positions of radar and point P at $t = 0$, and let the projection of the satellite move on to position B at time t , see Fig 3. In Fig 3 C is a point on the celestial sphere which has the same geocentric latitude as B and the same longitude as A (the position of the radar antenna focus at $t = 0$). The angle $B\hat{A}C = \nu_0$ and is the 'local heading angle' of the satellite at $t = 0$. The angle $A\hat{O}B$ is the angle through which the satellite has moved in its orbit from A to B and $A\hat{O}B = \Delta\alpha$. Also $B\hat{O}C = \Delta\Gamma$ and $A\hat{O}C = \Delta\Psi$. The change in longitude, $\Delta\Phi = B\hat{N}A$, $\Delta\Gamma$ and $\Delta\Phi$ are related by

$$\sin\frac{\Delta\Gamma}{2} = \sin\frac{\Delta\Phi}{2}\cos(\Psi + \Delta\Psi) \quad . \tag{46}$$

The analysis is somewhat intricate and for that reason a number of the intermediate steps are now set down.

The arc BC lies on a great circle and so ΔABC is a spherical triangle. The angle $A\hat{C}B$ is not exactly a right angle, although provided $\Delta\alpha$ is small (which it always is even for the 'spotlight' modes) then $A\hat{C}B$ is very nearly a rignt angle. For example $\Delta\alpha$ is usually of the order of 1 mrad for normal modes. Angle $A\hat{C}B$ is therefore written as $\pi/2 + \varepsilon$. From the sin formula for spherical triangles

$$\Delta\Gamma = \sin^{-1}\left(\frac{\sin \nu_0 \sin \Delta\alpha}{\cos \varepsilon}\right) \quad . \tag{47}$$

To find ε consider ΔCBN in which $N\hat{C}B = \pi/2 - \varepsilon$ and then

$$\varepsilon = \tan^{-1}\left[\sin(\Psi + \Delta\Psi)\,\frac{(1-\cos \Delta\Phi)}{\sin \Delta\Phi}\right] \quad . \tag{48}$$

Spherical triangle ABC gives

$$\Delta\Psi = \tan^{-1}\left[\tan \Delta\alpha \cos \nu_0 - \frac{\sin \Delta\Gamma \sin \varepsilon}{\cos \Delta\alpha \cos \Delta\Psi}\right] \tag{49}$$

which is an implicit formula for $\Delta\Psi$.

It will be evident that computing the coefficients in the range polynomial directly by means of the above formulae is not a practicable task. Fortunately, however, this is not necessary. All the angular changes $\Delta\psi$, $\Delta\Phi$, $\Delta\alpha$, $\Delta\Gamma$ and ε are of the same order of magnitude. Also $\Delta\alpha$ is a first order function of time. All terms in the range polynomial up to and including the third order term a_3 are required and hence it is only necessary to expand the expressions for $\Delta\gamma$, ε, $\Delta\Psi$ and $\Delta\Phi$ up to third order to obtain the required polynomial coefficients *exactly*. Expanding (46) gives:

$$\Delta\Phi = \Delta\Gamma \sec \Psi_0 + \Delta\Psi\Delta\Gamma \sec \Psi_0 \tan \Psi_0 + \Delta\Psi^2\Delta\Gamma \tan^2\Psi_0 \sec \Psi_0 + \frac{\Delta\Psi^2\Delta\Gamma}{2} \sec \Psi_0 + \frac{\Delta\Gamma^3}{24} \sec \Psi_0 \tan^2\Psi_0 + O(\Delta\Gamma\Delta\Psi^3,\ \Delta\Psi\Delta\Gamma^3) \tag{50}$$

(47) gives:

$$\Delta\Gamma = \Delta\alpha \sin \nu_0 - \frac{\Delta\alpha^3}{6} \sin \nu_0 \cos^2\nu_0 + \frac{\Delta\alpha\varepsilon^2}{2} \sin \nu_0 + O(\Delta\alpha^5, \Delta\alpha^3 t^2, \Delta\alpha\varepsilon^4) \tag{51}$$

(48) gives:

$$\varepsilon = \frac{\Delta\Phi}{2} \sin \Psi_0 + \frac{\Delta\Psi\Delta\Phi}{2} \cos \Psi_0 - \frac{\Delta\Psi^2\Delta\Phi}{4} \sin \Psi_0 + \frac{\Delta\Phi^3}{24} \sin \Psi_0 \cos^2\Psi_0 - O(\Delta\Psi\Delta\Phi^3, \Delta\Phi\Delta\Psi^3) \tag{52}$$

and (49) gives:

$$\Delta\Psi = \Delta\alpha \cos \nu_0 - \varepsilon\Delta\Gamma + \frac{\Delta\alpha^3}{3} \cos \nu_0 \sin^3\nu_0 + O(\Delta\alpha^2\Delta\Gamma\varepsilon, \Delta\Psi^2\Delta\Gamma\varepsilon, \Delta\Gamma^3\varepsilon, \Delta\Gamma\varepsilon^3) \quad .$$

..... (53)

These equations can be manipulated to give, finally:

$$\Delta\Psi = \Delta\alpha \cos \nu_0 - \frac{\Delta\alpha^2}{2} \sin^2\nu_0 \tan \Psi_0 + \frac{\Delta\alpha^3}{3} \cos \nu_0 \sin^3\nu_0 - \frac{\Delta\alpha^3}{2} \sin^2\nu_0 \cos \nu_0 \sec^2\Psi_0 + O(\Delta\alpha^4) \qquad (54)$$

$$\Delta\Phi = \Delta\alpha \sin \nu_0 \sec \Psi_0 + \Delta\alpha^2 \sin \nu_0 \cos \nu_0 \sec \Psi_0 \tan \Psi_0 - \frac{\Delta\alpha^3}{3} \sin \nu_0 \cos^2\nu_0 \sec \Psi_0 + \Delta\alpha^3 \sin \nu_0 \cos^2\nu_0 \tan^2\Psi_0 \sec \Psi_0 - \frac{\Delta\alpha^3}{3} \sin^3\nu_0 \tan^2\Psi_0 \sec \Psi_0 + O(\Delta\alpha^4) \quad , \qquad (55)$$

which are the required formulae for the changes in latitude and longitude in going from A to B in Fig 3, expressed in terms of change in satellite angle (or 'true anomaly') $\Delta\alpha$.

A problem arises here due to the singularity at the poles where $\Psi_0 = \pm\pi/2$ and it is necessary to avoid this case in the analysis. Thus the region within about 2° of the north and south poles is excluded. Fortunately this presents no difficulty because an exactly polar orbit is very unlikely to be used in practice. The suffix (0) on the variables in equations (54) and (55) refer to $t = 0$, *ie* point A in Fig 3.

Differentiating (54) and (55) and setting $t = 0$ gives:

$$\dot{\Psi}_0 = \dot{\alpha}_0 \cos \nu_0 \qquad (56)$$

$$\ddot{\Psi}_0 = \ddot{\alpha}_0 \cos \nu_0 - \dot{\alpha}_0^2 \sin^2\nu_0 \tan \Psi_0 \qquad (57)$$

$$\dddot{\Psi}_0 = \dddot{\alpha}_0 \cos \nu_0 - 3\ddot{\alpha}_0\dot{\alpha}_0 \sin^2\nu_0 \tan \Psi_0 + 2\dot{\alpha}_0^3 \cos \nu_0 \sin^3\nu_0 - 3\dot{\alpha}_0^3 \sin^2\nu_0 \cos \nu_0 \sec^2\Psi_0 \qquad (58)$$

$$\dot{\Phi}_0 = \dot{\alpha}_0 \sin \nu_0 \sec \Psi_0 \tag{59}$$

$$\ddot{\Phi}_0 = \ddot{\alpha}_0 \sin \nu_0 \sec \Psi_0 + 2\dot{\alpha}_0^2 \sin \nu_0 \cos \nu_0 \sec \Psi_0 \tan \Psi_0 \tag{60}$$

$$\begin{aligned}\dddot{\Phi}_0 = {} & \dddot{\alpha}_0 \sin \nu_0 \sec \Psi_0 + 6\ddot{\alpha}_0\dot{\alpha}_0 \sin \nu_0 \cos \nu_0 \sec \Psi_0 \tan \Psi_0 \\ & + 2\dot{\alpha}_0^3 \sin \nu_0 \cos^2 \nu_0 \sec \Psi_0 + 6\dot{\alpha}_0^3 \sin \nu_0 \cos^2 \nu_0 \tan^2 \Psi_0 \sec \Psi_0 \\ & - 2\dot{\alpha}_0^3 \sin^3 \nu_0 \tan^2 \Psi_0 \sec \Psi_0 \ .\end{aligned} \tag{61}$$

In addition to the motion of the satellite we also have to consider the rotation of the Earth. The latitude of a given point P remains constant but its longitude increases at a constant rate ξ rad/s. Thus

$$\dot{\phi} = \xi \ . \tag{62}$$

At this point the idea of the squint angle is introduced. The squint angle is here defined as the angle between the great circle arc AD and the great circle perpendicular to the orbit direction at A - see Fig 2. This definition is valid for general elliptic orbits as well as circular orbits since we consider the projection on to a sphere. If the squint angle is zero at the centre of the synthetic aperture at $t = 0$ then the radar is operating in the normal 'broadside' mode. This angle could be (and generally is) a function of time due to charges in the pitch and yaw angles of the spacecraft. The squint angle is defined to be positive when looking forwards.

Angle $\hat{NAD}$ (Fig 2) is then $\pi/2 - \nu_0 - \sigma_0$ where σ_0 is the squint angle at $t = 0$. Note that throughout this section ν_0 is defined as positive in Figs 2 and 3.

We are now in a position to compute the coefficients in the range polynomials.

4.3 The linear term

Differentiating (41), substituting (56) and (59) and simplifying by means of two spherical triangle formulae for triangle ADN, viz

$$\begin{aligned}\sin \psi_0 \cos \Psi_0 \cos(\nu_0 + \sigma_0) = {} & \sin \Psi_0 \cos(\Phi_0 - \phi_0) \cos \psi_0 \cos(\nu_0 + \sigma_0) \\ & + \sin(\Phi_0 - \phi_0) \sin(\nu_0 + \sigma_0) \cos \psi_0\end{aligned} \tag{63}$$

and

$$\sin\theta_0 \cos(\nu_0 + \sigma_0) = \cos\psi_0 \sin(\Phi_0 - \phi_0) \tag{64}$$

gives

$$\frac{d(\cos\theta_0)}{dt} = \dot{\alpha}_0 \sin\theta_0 \sin\sigma_0 + \xi \sin\theta_0 \cos\Psi_0 \cos(\nu_0 + \sigma_0) \tag{65}$$

which together with (43) gives:

$$a_1 = \frac{\dot{R}_0}{a_0}(R_0 - r\cos\theta_0) - \frac{R_0 r}{a_0}\dot{\alpha}_0 \sin\theta_0 \sin\sigma_0 - \frac{R_0 r}{a_0}\xi \sin\theta_0 \cos\Psi_0 \cos(\nu_0 + \sigma_0) \quad . \tag{66}$$

Thus the linear (range walk) term neatly separates into three terms. The first is zero for a circular orbit since $\dot{R}$ is then zero, and is thus associated with orbit eccentricity. The second term is zero when the squint angle σ_0 is zero and is thus associated with the radar squint. The third term is zero when ξ is zero and is thus associated with the rotation of the Earth. Thus linear range walk is in general caused by three factors; orbit eccentricity, squint and Earth rotation.

It is instructive to calculate a_1 for a typical example and the example chosen (Fig 4) is an image from Seasat Rev.762 over the east coast of England (the Wash and East Anglia). The coefficient a_1 will be calculated for the centre of the image. The Keplerian orbital elements for Seasat Rev.762 are given in Table 1:

Table 1

Orbital elements for Seasat Rev.762

Semi-major axis	7161.39494 km
Eccentricity	0.00186004
Inclination	108.0203°
Argument of perigee	148.1649°
Right ascension	89.3670°
Mean anomaly	252.3242°

R_0 and its derivatives and $\dot{\alpha}_0$ and higher derivatives can easily be worked out from the orbital elements using the usual orbit equations - see for example Ref 18, Chapter 3.

The latitude and longitudes of the satellite nadir and the centre of the image at $t = 0$ (the centre of the aperture) and values of r, R_0, $\dot{R}_0$, $\ddot{R}_0$, $\dot{\alpha}_0$, $\ddot{\alpha}_0$ and θ_0 are given in Table 2.

r is the Earth radius at P and is calculated on the basis of a spheroidal Earth with a flattening factor of 1/298.3 and equatorial radius of 6.37816×10^6 m.

Table 2

Parameters of centre of Fig 4 at t = 0

Latitude of nadir	51.25^o
Longitude of nadir	-2.08^o
Latitude of point	52.78^o
Longitude of point	0.67^o
θ_0	2.265^o
r	6.36444×10^6 m
R_0	7.16297×10^6 m
$\dot{R}_0$	1.37760×10 m/s
$\ddot{R}_0$	-1.758803×10^{-3} m/s^2
$\dot{\alpha}_0$	1.041312×10^{-3} rad/s
$\ddot{\alpha}_0$	4.005×10^{-9} rad/s^2
ξ	0.7292115×10^{-4} rad/s
ν_0	29.20^o

The corresponding values of the three terms in the expression for a_1 are:

Eccentric term: $$\frac{\dot{R}_0}{a_0}(R_0 - r\cos\theta_0) = 13.147 \text{ m/s}$$

Squint term: $$-\frac{R_0 r}{a_0}\dot{\alpha}_0 \sin\theta_0 = -38.887 \text{ m/s per degree of squint for small squint angles}$$

Rotation term: $$-\frac{R_0 r}{a_0}\xi \sin\theta_0 \cos\Psi_0 \cos\nu_0 = -82.385 \text{ m/s}.$$

The pitch and yaw angles of the satellite of the centre of the aperture were about -0.0040 rad and 0.0112 rad. Because of the steep (low) incidence angle

the pitch angle has a greater effect on the squint than the yaw and in this instance they nearly cancel out to give a squint angle of about -0.092°. Hence the squint term amounts to some 3.58 m/s. The squint angle can easily be 0.5° or so on occasions for SEASAT.

The linear (range walk) term is then about 65.66 m/s for this example.

One can draw some very interesting conclusions from equation (66). Notice that it is possible to select a squint angle σ_0 such that a_1 is zero. Hence linear range walk could, in principle, be eliminated for a specific value of θ_0 corresponding to, say, the middle of the swath by varying the squint angle of the radar. Moreover if the orbit is circular then

$$a_1 = -\frac{R_0 r \sin\theta_0}{a_0}\left[\dot{\alpha}_0 \sin\sigma_0 + \xi \cos\Psi_0 \cos(\nu_0 + \sigma_0)\right] \qquad (67)$$

and a_1 would be zero over the *whole swath width*.

In practice the centre of the Doppler spectrum (which is amplitude modulated by the antenna beam pattern) would be measured on board the satellite and the satellite yaw or (better) the pitch angle would be varied via the satellite attitude control system until the Doppler spectrum is centred on zero. This type of technique could have far reaching implications for the design of a SAR and processor system for real-time processing. The range walk problem is explained further in section 12.

4.4 The quadratic term

Differentiating (41) twice, substituting from (56), (57), (59) and (60) and simplifying by means of (41) and an additional formula for triangle ADN:

$$\sin\theta_0 \sin(\nu_0 + \sigma_0) = \sin\psi_0 \cos\Psi_0 - \cos\psi_0 \sin\Psi_0 \cos(\Phi_0 - \phi_0) \qquad (68)$$

gives

$$\begin{aligned}\frac{d^2}{dt^2}(\cos\theta_0) = &-\dot{\alpha}_0^2 \cos\theta_0 + \ddot{\alpha}_0 \sin\theta_0 \sin\sigma_0 \\ &+ 2\xi\dot{\alpha}_0\left[\sin\nu_0 \cos\psi_0 \cos(\Phi_0 - \phi_0) - \cos\nu_0 \cos\psi_0 \sin\Psi_0 \sin(\Phi_0 - \phi_0)\right] \\ &- \xi^2 \cos\Psi_0 \cos\psi_0 \cos(\Phi_0 - \phi_0) \quad . \qquad (69)\end{aligned}$$

The terms including ξ can be simplified further but only by introducing yet another angle and so it has been left as it is. In conjunction with (43) and (45), (69) gives:

$$a_2 = \frac{1}{2a_0}\left[\ddot{R}_0 + \frac{2\dot{R}_0^2}{R_0}\right]\left[R_0 - r\cos\theta_0\right] + \frac{1}{2a_0}\left[\dot{R}_0^2 - a_1^2\right] - \frac{\dot{R}_0 a_1}{R_0}$$

$$- \frac{R_0 r}{2a_0}\Big[-\dot{\alpha}_0^2\cos\theta_0 + \ddot{\alpha}_0\sin\theta_0\sin\sigma_0$$

$$+ 2\xi\dot{\alpha}_0\Big\{\sin\nu_0\cos\psi_0\cos(\Phi_0 - \phi_0) - \cos\nu_0\cos\psi_0\sin\Psi_0\sin(\Phi_0 - \phi_0)\Big\}$$

$$- \xi^2\cos\Psi_0\cos\psi_0\cos(\Phi_0 - \phi_0)\Big] \qquad (70)$$

which is a general expression for the quadratic (range curvature) term for elliptic orbits and rotating Earth. It can be separated into terms associated with eccentricity (the terms containing derivatives of R and second derivative of α), and those associated with Earth rotation (the terms containing ξ). For a circular orbit, stationary Earth and zero squint angle the expression simplifies drastically to give

$$a_2' = \frac{R_0 r\dot{\alpha}_0^2\cos\theta_0}{2a_0} = \frac{V_p^2}{2a_0}\left(1 - \frac{h}{R_0}\right)\cos\theta_0 \qquad (71)$$

where V_p is the satellite velocity $= R_0\dot{\alpha}_0$ and h is the height of the satellite above the Earth. Since θ_0 is usually small (2° to 3°) and $h \simeq R_0/10$ then the usual simple expression given in the literature for a_2, $V_p^2/2a_0$, see for example Ref 9, equation (1), is correct to within 10% or so. It is again instructive to determine the magnitudes of the various terms in (70) for the 'typical example' Fig 4. Substituting values from Table 2 into the part of (70) which contains the radial derivatives gives:

$$\frac{1}{2a_0}\left[\ddot{R}_0 + \frac{2\dot{R}_0^2}{a_0}\right]\left[R_0 - r\cos\theta_0\right] + \frac{1}{2a_0}\left[\dot{R}_0^2 - a_1^2\right] - \frac{\dot{R}_0 a_1}{R_0} = -3.795 \times 10^{-3}\ \mathrm{m/s^2}\ .$$

..... (72)

The remainder of (70) contains the angular derivatives:

$$- \frac{R_0 r}{2a_0}\Big[- \dot{\alpha}_0^2 \cos\theta_0 + \ddot{\alpha}_0 \sin\theta_0 \sin\sigma_0$$

$$+ 2\xi\dot{\alpha}_0\Big\{\sin\nu_0 \cos\psi_0 \cos(\Phi_0 - \phi_0) - \cos\nu_0 \cos\psi_0 \sin\Psi_0 \sin(\Phi_0 - \phi_0)\Big\}$$

$$- \xi^2 \cos\Psi_0 \cos\psi_0 \cos(\Phi_0 - \phi_0)\Big] = 28.0946 \text{ m/s}^2 \quad . \qquad (73)$$

If the maximum synthetic aperture time is taken at ±1 s, and if the maximum tolerable phase error over the aperture is taken to be equivalent to $\lambda/16$ where λ is the wavelength, then it is obvious that the terms containing the radial derivatives are negligible. The term containing $\ddot{\alpha}_0$ is also negligible ($R_0 r\ddot{\alpha}_0 \sin\theta_0 \sin\sigma_0/2a_0 = 4.86 \times 10^{-5}$ m/s^2).

There is no doubt that the terms associated with eccentricity can be ignored in this instance and that a_2 can be calculated on the basis of a circular orbit. Hence such phenomena as the precession of the perigee and variation of eccentricity have no material effect on the quadratic term - provided that the orbit eccentricity is small enough. A more refined analysis can set limits on the orbit eccentricity but this would take us too far from the main theme of this paper.

4.5 The cubic term

The terms involving eccentricity in the quadratic case were shown to be negligible and they are even more so in the cubic case. Unfortunately all the other terms (of which there are very many) in the equation for a_3 only differ by one order of magnitude or so and all must therefore be considered. The resulting expression for a_3 will not be set down here; it can easily be calculated along the same lines as for a_1 and a_2. The expressions are complicated and not easily simplified due to the inherent asymmetry. However if this is done it will be found that substituting the values in Table 2 for the example given before gives $a_3 = 1.26 \times 10^{-2}$ m/s^3. Hence over a synthetic aperture of ±1 s there is a variation of about 1/9 wavelength. This, unfortunately, is on the limits of the maximum acceptable error, and to be on the safe side in this case the cubic term should be included. The example is not, of course, a worst case example and the cubic term increases at more northerly latitudes.

For SAR systems with shorter wavelengths (such as the proposed ERS-1 SAR with $\lambda \simeq 5.6$ cm) and hence shorter synthetic aperture intervals, the cubic term will be completely negligible. Conversely with systems which have a long

wavelength and long synthetic aperture interval (such as the proposed spotlight mode for the Shuttle SIR-B mission) the cubic term will be significant.

4.6 The range polynomial: conclusion

In this section it has been shown how it is possible to calculate the terms in the range polynomial for the most general case of an elliptical orbit and rotating Earth. It has been shown that the eccentricity of the radar platform orbit has most effect on the linear term, for small eccentricity it can be neglected so far as the quadratic and cubic terms are concerned. A more refined analysis is possible based on this section but will not be presented here. One implication of this is that the use of nominal orbital parameters is more likely to have an effect on the linear range walk correction in the processing than on focussing via the quadratic phase term. The use of 'auto-focussing' techniques to overcome the effects of the use of nominal orbital elements is thus somewhat questionable. All auto-focussing techniques known to the author rely in some way on the defocussing being caused by an error on the quadratic phase over the correlation. An error on the linear range walk can also cause defocussing due to the selection of incorrect data in the range walk correction procedure and this is discussed again in section 13.

5 DIGITAL PROCESSING: MATHEMATICAL PRELIMINARIES

In this section some relevant background theory is briefly presented. Most of this theory will be found in one form or another in Ref 16. Consider a function of time $f(t)$; $f(t)$ may be complex valued. Its Fourier transform $F(\omega)$ is

$$F(\omega) = \int_{-\infty}^{\infty} f(t) \exp - i\omega t \, dt \quad . \tag{74}$$

The inverse transform is

$$F(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) \exp i\omega t \, d\omega \quad . \tag{75}$$

$F(\omega)$ is the spectrum of the function $f(t)$, and is a complex valued function. If $f(t)$ is purely real or imaginary then the real and imaginary parts of $F(\omega)$ have symmetry properties:

if $f(t)$ is real valued then $F^*(\omega) = F(-\omega)$ (76)

if $f(t)$ is imaginary valued then $F^*(\omega) = -F(-\omega)$ (77)

where the asterisk indicates complex conjugate.

5.1 The analytic signal

From equations (76) and (77) we see that a real (or imaginary) function can therefore be completely determined by only the positive (or negative) frequencies, *ie* only half of the spectrum because the other half can be reconstructed via equations (76) and (77). This takes us to the idea of the analytic signal.

Let $f(t)$ be purely real with $F(\omega)$ given by (74). Then $F^*(\omega) = F(-\omega)$. Let $g(t)$ be purely imaginary with Fourier transform $G(\omega)$. Then $G^*(\omega) = -G(-\omega)$.

Let $g(t)$ be such that for $\omega > 0$, $G(\omega) = F(\omega)$.

Then $G^*(\omega) = F^*(\omega)$ as well, for $\omega > 0$.

Then for $\omega < 0$ we have $G(-\omega) = -F(\omega)$.

Let $z(t) = f(t) + g(t)$, then its Fourier transform is

$$\begin{aligned} F(\omega) + G(\omega) &= 2F(\omega) && \omega > 0 \\ &= 0 && \omega < 0 \end{aligned}$$

and we therefore have a single sided spectrum. Usually $g(t)$ is written as $i\hat{f}(t)$ and then

$$\begin{aligned} z(t) &= f(t) + i\hat{f}(t) \\ z(\omega) &= 2F(\omega) && \omega > 0 \\ &= 0 && \omega < 0 \ . \end{aligned}$$

And so starting with a real function $f(t)$ we can form its Fourier transform $F(\omega)$, double it, and set its negative frequencies equal to zero. If this function is then inverse transformed, the real part of the resulting complex valued function $z(t)$ is equal to the original real function $f(t)$. Thus only half of the spectrum of the real valued $f(t)$ is needed to define it completely.

Now consider the function

$$a(t) \exp i\phi(t) = a(t) \cos \phi(t) + ia(t) \sin \phi(t)$$

where $a(t)$ and $\phi(t)$ are real valued. $a(t) \exp i\phi(t)$ is assumed to be analytic (and is an entire function of exponential type with singularities only at infinity). Then it can be proved (see Ref 20, Chapter 5) that the function has a single sided spectrum. Hence it is possible to identify $a(t) \cos \phi(t)$ with $f(t)$ and $a(t) \sin \phi(t)$ with $\hat{f}(t)$. In actual fact $f(t)$ and $\hat{f}(t)$

are mutual Hilbert transforms, but that is only of passing interest here. It will be observed that the chirp signal, equation (3), is in this form and is an analytic signal. Likewise (1) is the analytic signal corresponding to the real signal (2).

Thus if we have a real valued function such as $a(t) \cos \phi(t)$ writing it as $a(t) \exp i\phi(t)$ gives the 'analytic signal' with a single sided spectrum.

It will be shown in section 6.1 that, in the case of a SAR system in which the Doppler band extends over both positive and negative frequencies, it is absolutely necessary to operate on analytic signals.

5.2 The convolution and correlation theorems

If

$$g(t) = \int_{-\infty}^{\infty} f(\tau)h(t - \tau)d\tau = \int_{-\infty}^{\infty} f(t - \tau)h(\tau)d\tau \tag{78}$$

then

$$G(\omega) = F(\omega)H(\omega) \tag{79}$$

(see Ref 16, p.21, equation (1-47)). The convolution is usually written

$$g(t) = f(t) \otimes h(t) \quad . \tag{80}$$

Hence convolution in the time domain corresponds to multiplication in the frequency domain (and vice versa).

A similar theorem exists for the correlation function:

$$\text{If} \qquad c(t) = \int_{-\infty}^{\infty} f^*(\tau)h(t + \tau)d\tau = \int_{-\infty}^{\infty} f^*(\tau - t)h(\tau)d\tau \tag{81}$$

then

$$C(\omega) = F^*(\omega)H(\omega) \quad . \tag{82}$$

In Ref 21, equation (406) p.768, it is shown that the complex response from a matched filter is determined by a correlation process.

To perform matched filtering in the frequency domain therefore, one Fourier transforms the signal and multiplies it by the complex conjugate of the chirp replica spectrum and then performs the inverse transformation. This is the basis of frequency domain compression in both the range and azimuth directions.

5.3 The sampled discrete Fourier transform

The discrete Fourier transform pair analogous to the continuous transforms are:

$$F(\ell) = \frac{1}{N} \sum_{k=0}^{N-1} f(k) \exp - \frac{i2\pi k\ell}{N} \tag{83}$$

$$f(k) = \sum_{\ell=0}^{N-1} F(\ell) \exp \frac{i2\pi k\ell}{N} \quad . \tag{84}$$

This pair of transforms can be regarded as a mapping of the sequence $f(k)$, $k = 0 \ldots N-1$ on to the sequence $F(\ell)$, $\ell = 0 \ldots N-1$ and vice versa. There are N samples in the 'time' domain and N in the 'frequency' domain.

The reader is reminded that it is implicit in this representation that the original continuous time function $f(t)$ and the corresponding spectrum $F(\omega)$ have both been periodically repeated. Usually the temporal function $f(t)$ has a finite length and is sampled at intervals of Δt, say, so that if T is the total length $T = N\Delta t$. The corresponding spectrum, however, is not finite and the process of periodically repeating it causes the edges to be periodically folded into the fundamental period. This is called aliasing. The width of the fundamental frequency period is $1/\Delta t$, see Ref 22, and it is therefore necessary that there shall be no frequencies higher than $f_s/2$, where $f_s = 1/\Delta t$, present in the original spectrum. The reason for the additional factor of $\frac{1}{2}$ is that we have both the positive and negative frequencies to contend with (see Fig 5). The frequency $f_s/2$ is commonly called the folding frequency. It will be observed that the sample sequence $F(\ell)$ runs from 0 to $N-1$ and therefore in the first half of the sampled spectrum we have positive frequencies and in the second half negative frequencies. The spectrum is folded about $f_s/2$.

There are no aliasing problems with the time sampled sequence because it is a finite length record and is zero outside the sampled range. Each sample in the time domain is separated by Δt where $\Delta t = T/N$ and each sample in the frequency domain is separated by $\Delta f = 1/T$ where $\Delta f = F/N$ and $F = f_s$ the sampling frequency.

This discussion applies to a real valued function. When the sampled function is complex valued with a single sided spectrum then the spectrum can be periodically repeated at double the rate of a double sided spectrum without ambiguity. This means that if the sampling frequency is f_s the spectrum can have frequencies up to f_s instead of $f_s/2$ before ambiguity sets in. This can be

illustrated by a simple example. Consider the real and complex signals $\cos 2\pi ft$ and $\exp i2\pi ft$. Suppose both the signals are sampled at times n/f_s, $n = 0 \ldots N-1$. Then for the complex signal frequencies f and $f + f_s$ are ambiguous because

$$\exp \frac{i2\pi(f+f_s)n}{f_0} = \exp \frac{i2\pi f n}{f_0} \tag{85}$$

and for the real signal the frequencies f and $f + f_s/2$ are (just) ambiguous because

$$\cos\left[2\pi \frac{(f+f_s/2)}{f_s}\right] = -\cos \frac{2\pi f}{f_s} \quad . \tag{86}$$

If real signals with frequencies in the range 0 to $f_s/2$ are sampled at f_s then the complex signal may be sampled at $f_s/2$ without ambiguity.

5.4 The discrete convolution and correlation theorems

If

$$g(k) = \frac{1}{N}\sum_{m=0}^{N-1} f(m)h(k-m) = \frac{1}{N}\sum_{m=0}^{N-1} f(k-m)h(m) \tag{87}$$

then

$$G(\ell) = F(\ell)H(\ell) \qquad \ell = 0 \ldots N-1. \tag{88}$$

The convolutions in equation (87) and (88) are cyclic because $f(m)$ and $h(m)$ $m = 0 \ldots N-1$ are periodic functions. Again, then, convolution in the sampled time space corresponds to multiplication in the sampled spectrum space.

There is a corresponding result for the correlation.

$$\text{If} \qquad c(k) = \frac{1}{N}\sum_{m=0}^{N-1} f^*(m)h(k+m) = \frac{1}{N}\sum_{m=0}^{N-1} f^*(m-k)h(m) \tag{89}$$

then

$$c(\ell) = F^*(\ell)H(\ell) \quad , \tag{90}$$

again, this correlation is cyclic.

Cyclic convolutions can produce problems (see Ref 22, p.48). For example, in the case of the range compression process it will be demonstrated that there

are end effects which make it necessary to process more data than is required by the image.

5.5 Interpolation via the Fourier transform

In terms of its discrete Fourier transform (DFT) a sampled function is:

$$f(k\Delta t) = \sum_{\ell=0}^{N-1} F(\ell\Delta\omega)\, \exp \frac{i2\pi k\ell}{N} \qquad \left.\begin{matrix} 0 \leqslant k \leqslant N-1 \\ 0 \leqslant \ell \leqslant N-1, \end{matrix}\right\} \quad (91)$$

(see equation (84)). Δt is the spacing of the time samples and $\Delta\omega$ is the spacing of the angular frequency samples. $\Delta\omega = 2\pi\Delta f$ where Δf is the frequency sample spacing

$$\Delta t = \frac{1}{N\Delta f} = \frac{2\pi}{N\Delta\omega} \quad . \qquad (92)$$

Suppose that the length of the inverse DFT is doubled and define a new set of spectrum samples $F'(\ell\Delta\omega)$ such that

$$\begin{aligned} F'(\ell\Delta\omega) &= F(\ell\Delta\omega) && \text{for } 0 \leqslant \ell \leqslant N-1 \\ &= 0 && \text{for } N \leqslant \ell \leqslant 2N-1. \end{aligned} \qquad (93)$$

Then the new inverse DFT is:

$$\sum_{\ell=0}^{2N-1} F'(\ell\Delta\omega)\, \exp \frac{i2\pi\ell k}{2N} = \sum_{\ell=0}^{N-1} F(\ell\Delta\omega)\, \exp \frac{i2\pi\ell k}{2N}$$

$$= \sum_{\ell=0}^{N-1} F(\ell\Delta\omega)\, \exp \frac{i2\pi\ell(k/2)}{N} \quad . \qquad (94)$$

Comparing this with equation (91) gives:

$$\sum_{\ell=0}^{2N-1} F'(\ell\Delta\omega)\, \exp \frac{i2\pi\ell k}{2N} = f\,\frac{k\Delta t}{2} \quad . \qquad (95)$$

Since the number of samples in the inverse DFT was doubled, the numbers of samples of the time function $f(t)$ was also doubled and there are therefore

2N samples at intervals of $\Delta t/2$. That is to say the time function has been interpolated. All the original samples exist and in addition there are new (interpolated) samples midway between the original samples.

This process can evidently be continued and if the inverse DFT is performed with a length 2^n times the forward DFT this gives $2^n - 1$ in erpolation points.

Some further comments will be made on the process in section 6.3. There is another interpolation technique using the DFT which is not limited to dividing the sample interval by some power of 2. Returning to equation (91) it is evident that

$$f(k\Delta f + \alpha\Delta t) = \sum_{\ell=0}^{N-1} F(\ell\Delta\omega) \exp \frac{i2\pi k\ell}{N} \exp \frac{i2\pi\alpha\ell}{N} \quad . \qquad (96)$$

It is therefore possible to produce new samples at any chosen point within the original intervals ($\alpha < 1$) simply by weighting the spectrum samples by a phase shift $\exp(i2\pi\alpha\ell/N)$ before inverse Fourier transforming.

Both of the above techniques have successfully been used on SAR processing at RAE.

6 SPECTRA

6.1 The range spectrum

The signal transmitted by a SAR is usually a frequency modulated 'chirp' with a large time-bandwidth product. The spectrum of such a chirp is calculated in Ref 16 (example 8-5, p.270). It is a Fresnel integral. If the chirp is centred on f_0 (the radar 'carrier' or centre frequency) and has a nominal width of ΔF (= the chirp rate in Hz/s × the pulse width) then the spectrum is as sketched in Fig 6. This is the spectrum of the real valued signal (2) as obtained by the radar receiver from a point scatterer. The actual signal as it exists in space is complex valued and has a single sided spectrum for reasons which will not be pursued here.

Fig 2 shows the modulus of the complex valued spectrum as do all the subsequent figures.

Each point on the ground reflects a replica of the transmitted chirp and the received signal at any instant is the linear sum of chirps reflected from points illuminated by the footprint and within a distance corresponding to half a pulse width. The bandwidth of the received signal is therefore the same as

that of a single chirp, although, of course, the detail of the spectrum is greatly different and is noise-like being composed of chirps with random phases and displaced in time.

The actual received spectrum is sketched in Fig 7.

The receiver coherently mixes the signal down to the offset video frequency. This offset is chosen to be slightly greater than $\Delta F/2$ to avoid aliasing the spectrum and is one quarter of the subsequent digitising frequency for a real valued signal. The resulting spectrum is shown in Fig 8.

The signal is still real valued at this point. At this stage the signal could be operated on by a quadrature filter to give the analytic signal in which case the negative half of the spectrum disappears and then f_1 could be mixed down to zero to give a spectrum symmetric about zero. The resulting signal would be complex valued. For the moment, however a real valued signal will still be used.

The signal is next digitised at a sampling frequency f_s (typically exactly four times the offset video f_1). The process of sampling the signal causes the spectrum to be periodically repeated at a frequency of f_s - see Fig 9. Just the envelope of the spectrum is shown in Fig 9, the positive and negative frequencies are indicated by '+' and '-'.

When range compression is performed in the frequency domain (as it nearly always is) using a discrete Fourier transform (the 'fast' Fourier transform version) then the spectrum is also sampled. If N samples are used in the Fourier transform the spectrum is sampled with a frequency increment of f_s/N from sample to sample, see Fig 10. In the case of the fast Fourier transform there will be 2^n samples, typically 4096, 8192 or 16384 for a spacecraft SAR. Notice that the positive frequencies are contained in samples 1 to $N/2$ and the negative frequencies in samples $N/2 + 1$ to N. The chirp spectrum is similar and thus the product of the chirp and signal spectra is similar.

At this stage the negative frequencies are eliminated by setting samples $N/2 + 1$ to N equal to zero, and on inverse transforming, the compressed analytic signal is produced.

6.2 The azimuth spectrum

Previously the spectrum of just one pulse was considered. Now, the spectrum of a large number of pulses is discussed. Fig 8 gives the spectrum of just one received pulse at the offset video frequency. The radar periodically repeats pulses at a pulse recurrence frequency of ν Hz, say. The effect of this is to turn the spectrum in Fig 8 into a line spectrum. The distance between the

lines is ν Hz . The resulting spectrum of a large number of received pulses is shown in Fig 11. This should not be confused with the line spectrum already considered in Fig 10. That spectrum arose in the range compression process, a process independent of the azimuth direction.

The spectrum shown in Fig 11 would be obtained if the radar platform were stationary. The platform moves, of course, and each of the lines in the spectrum therefore moves due to the Doppler effect.

Suppose that range compression is performed using real data. The result is given in equation (6). The Doppler 'phase' term is a cosine function. If range compression is performed with an analytic signal the result is given in equation (7) and the phase term is an exponential. This is a crucial difference. Consider just one line in the spectrum of Fig 11. If the real 'phase' term is used then a Doppler shift splits the line into two lines because the cosine term implicitly includes both positive and negative frequencies - see Fig 12. In other words positive and negative frequencies are indistinguishable, *ie* aliassed. For the analytic signal with a single sided spectrum a Doppler shift just shifts each line as shown in Fig 13. There is only one line and positive and negative frequencies are kept separate.

In actual fact one has a continuum of Doppler frequencies both positive and negative for a broadside looking radar. Each line is then broadened into a continuum - see Fig 14.

Note that the radar is designed in such a way that the pulse recurrence frequency is high enough to sample the Doppler band defined by the antenna beam-width and this is done on the basis of an analytic signal.

If the radar is a pure squint mode radar with only positive or negative frequencies then the aliassing problem would not exist. One could then either choose a high enough pulse recurrence frequency or a very carefully chosen prf so that the Doppler band is aliassed into a wholly positive or negative band. This would undoubtedly lead to problems both with the design of the radar and the spacecraft attitude control.

In any event for a broadside looking radar the analytic signal is always an absolute necessity and is probably also necessary for practical reasons for a pure squint mode radar.

6.3 Interpolation

In section 5.5 an interpolation technique was described in which the data to be interpolated are Fourier transformed via a DFT and then inverse transformed via a DFT with a larger number of samples. In using this technique one has to be

careful to treat the spectrum correctly. This will now be described, with reference to interpolating range compressed data.

Consider the sampled range spectrum shown in Fig 8. It corresponds to a Fourier transformed pulse. Both pulse and spectrum are sampled and periodically repeated. The period of the spectrum is f_s, the sampling frequency. Suppose that there are N samples. $N/2$ cover the negative frequencies and $N/2$ the positive frequencies, see Fig 9. Increasing the number of samples in the inverse transform is equivalent to sampling at a higher rate and the spectrum is periodically repeated at a longer period mf_s, say, where m is an integer equal to some power of 2. The resulting spectrum is shown in Fig 15. In order to achieve this the negative frequencies which occupy samples $N/2 + 1$ to N in the original spectrum must be moved up to samples $mN - N/2 + 1$ and mN in the new spectrum before inverse transforming as shown in Fig 15. Everything will then be correct and the correct interpolated samples will be produced.

7 SIDELOBE REDUCTION - APERTURE AND SPECTRUM WEIGHTING

Equation (15) gives the point spread function in both range and azimuth for the ideal matched filtering process. The modulus of the point spread function in both directions is of the form $|\sin x/x|$. This function is plotted in Fig 16. It will be observed that together with the main lobe there are sidelobes. These sidelobes are usually considered to be objectionable, although the author has processed images both with sidelobe reduction and without and can rarely tell the difference. The sidelobes are noticeable in the case of very bright point-like scatterers.

It is possible to reduce the level of these sidelobes but only at the expense of broadening the main lobe, *ie* reducing the resolution.

If the correlation process (matched filtering) is performed in the time domain then an integral of the form:

$$|g(\tau)| = \left| \int_{t_0-T/2}^{t_0+T/2} \exp i2\pi\tau\Delta F \frac{t}{T} \, dt \right| \tag{97}$$

is obtained. If the correlation process is performed in the frequency domain an integral of a similar form is obtained (this is the inverse Fourier transform of the product of the spectra of the signal and chirp replica or phase history replica).

$$|g(\tau)| = \left| \int_{f_0-F/2}^{f_0+F/2} \exp\, i2\pi\tau\Delta F\, \frac{f}{F}\, df \right| \qquad (98)$$

where the time centre of the aperture is t_0, the frequency centre is f_0, the time width is T, and the frequency width is F (this is equal to ΔF but formula (98) has been written as it has in order to exhibit symmetry with (97).

Sidelobe reduction is obtained by weighting the integrals in (97) and (42). A discussion will be found in Ref 21, section 3.4.2 p.780. Such weighting functions are usually expressed as a Fourier series with a period of either T or F. There are very many of them. The ideal function (ideal in the sense that one obtains minimum mainlobe broadening for a given sidelobe reduction) is the Dolph-Chebycheff function - see Ref 21, section 3.4.2.2 p.782. In Ref 16 Papoulis lists a number of different functions (section 7.3, p.234).

By far the most popular weighting function is a 'raised cosine'. This will be recognised as simply the first two terms in a truncated Fourier series. This is also known as a Taylor weighting function - see Ref 21.

The time weighting function is:

$$w(t) = 1 + 2\beta \cos\left[\frac{\pi(t-t_0)}{T/2}\right] \qquad t_0 - T/2 \leqslant t \leqslant t_0 + T/2 \qquad (99)$$

and the frequency weighting version is:

$$w(f) = 1 + 2\beta \cos\left[\frac{\pi(f-f_0)}{\Delta F/2}\right] \qquad f_0 - \Delta F/2 \leqslant f \leqslant f_0 + \Delta F/2 \qquad (100)$$

β is a parameter $0 \leqslant \beta \leqslant \frac{1}{2}$ which specifies the sidelobe levels. The resulting point spread function is:

$$|g(\tau)| = \frac{\sin \pi\Delta T\tau}{\pi\Delta F\tau}\left[1 + \frac{2\beta}{(\tau\Delta F)^{-2} - 1}\right] \qquad (101)$$

and the phase part of the psf is unaltered.

The signal energy is increased by a factor of $1 + 2\beta^2$ in both frequency and time domain correlations since

$$\int_{-T/2}^{T/2} w^2(t)dt = 1 + 2\beta^2 \ . \qquad (102)$$

The first zero in the point spread function is located at $\sqrt{1/(1-2\beta)}/\Delta F$ the position of the other zeros are unchanged. The relationship between β and the sidelobe levels must be obtained numerically.

An alternative to the Taylor weighting function is the Chebyshev weighting function, as pointed out above. This weighting function is best designed by means of the Remez exchange algorithm - see Ref 23, p.136. The sidelobe levels are specified and the number of weights, and then the design algorithm gives the weights. Sidelobe weighting using this type of weighting function has been used in Ref 9.

8 RANGE COMPRESSION

The various aspects of the range compression process have been discussed in previous sections and now in this section they are collected together.

Range compression is usually performed in the frequency domain using the correlation theorem in section 5.2 and the discrete Fourier transform in section 5.3. The fast Fourier transform (FFT) version of the discrete Fourier transform (DFT) is invariably employed - see Ref 22 for a description of the FFT. The FFT is just an efficient way of computing a DFT. In order to use it the number of samples must be some power of 2, *ie* $N = 2^n$.

The question of how many pulses and which pulses to compress is considered in the sections on prefiltering and azimuth compression.

It is interesting to examine the difference between correlation via the time domain and via the FFT. The number of multiplications and additions for the FFT is given in Ref 22. If there are N complex samples in the signal and n complex non-zero samples in the replica then correlation via the FFT requires approximately $N(1+\log_2 N)$ multiplications and $2N \log_2 N$ additions, and straight correlation requires nN multiplications and nN additions. $N > n$ because the replica is much shorter than the signal for obvious reasons.

As an example consider SEASAT. If one chose 8192 real range samples out of a total of 13680, *ie* 4096 complex samples for the analytic signal, and 1536 real value samples (768 complex) for the chirp replica then the FFT correlation requires 53248 multiplications and 98304 additions and subtractions. Straight correlation requires over 3 million multiplications and 3 million additions. Correlation via the FFT is therefore much faster.

All pulses are correlated against the same chirp replica. This chirp replica is known from the system specifications. Sometimes it has to be adjusted slightly in order to achieve the best possible range focussing.

The chirp replica, then, is sampled at the same rate as the data. The chirp record length is packed out with zeros to make it the same length as each pulse record length (= 2^n, as pointed out above). The packed record is then Fourier transformed and conjugated, the resulting spectrum is then weighted by the side-lobe weighting function given in section 7, suitably sampled. Naturally the spectral width of the weighting function ΔF is the same as that of the chirp spectrum. The weighted chirp spectrum is then permanently stored and used to correlate all the pulses. Each pulse signal is taken out of bulk storage and either packed with zeros to 2^n or truncated (this is more usual). Each pulse is then Fourier transformed, multiplied by the weighted chirp spectrum, and inverse Fourier transformed. This process is continued until all the pulses required to process the image have been compressed.

The reason for the truncation (or packing) of the pulse samples is that the total number of samples in each pulse is never, in practice, an exact power of 2. For example in the case of SEASAT N = 13680. There are then two choices and, usually, considerations of total storage for the range compressed pulses dictate that each pulse be truncated to 8192 or less.

It will be recalled that the FFT correlation is a cyclic correlation - see section 5.4, and this will now be examined. The signal samples can (schematically) be placed at equiangular points around a circle, and the chirp replica also, see Fig 17. The cyclic convolution then implicitly involves multiplying each signal sample with the corresponding replica sample, adding, and then rotating the replica relative to the signal by one sample and repeating the process N times. The FFT correlation does not do this explicitly, of course, but is entirely equivalent.

If the signal has been truncated it is easily seen that some signal has been lost for a time of $T/2$ at each truncation point, where T is the chirp width. It will be recalled, that due to the finite chirp width, the signal reflected from each point scatter does not extend over the whole signal but only over an interval of T. Hence if both the beginning and end of each pulse have been truncated the signals will not be properly correlated over an interval of $\pm T/2$ around the 'join' in the cyclic correlation. The top and bottom of the resulting image will then be spoilt over a distance corresponding to half a pulse width. It is therefore necessary to bear this in mind when selecting range samples to cover a particular image and a contingency of $T/2$ is required at the beginning and end of each pulse. An additional contingency is required to cover range migration and this is explained further in section 13.

It remains to be shown how the offset video frequency f_1 is removed, see equation (7) and subsequent remarks. This can be done immediately before inverse Fourier transforming each product spectrum, *ie* after each signal spectrum has been multiplied by the replica spectrum. As pointed out in section 2.1 the offset video is often arranged to be one quarter of the real sampling frequency. This is one half of the complex sampling frequency (the product spectrum at this stage is that of the analytic signal). If the offset video f_1 is exactly half of the maximum frequency in the spectrum then to translate it to zero the spectrum is simply moved by one half period. The procedure, then, is to rotate the product spectrum in either direction by $N/2$ before inverse Fourier transforming.

The offset video can also be removed in the time domain, after inverse transforming, by reversing the sign of alternate samples. Equation (7) gives the phase progression due to the offset video as $\exp i2\pi f_1 t_n$. If f_0 is the sampling frequency of the real data, $f_s/2$ is the sampling frequency of the complex analytic signal and the time increment from sample to sample is $2/f_s$. Hence $t_n = n(2/f_s)$ and $f_1 = f_s/4$ and then $\exp i2\pi f_1 t_n = \exp i\pi n = (-1)^n$. These procedures can obviously be generalised to cope with the situation in which the offset video frequency is not exactly one quarter of the sampling frequency.

9 LOOK FILTERING

In the processing of an image the Doppler spectrum is usually divided into a number of equal parts and a separate image is processed from each part. Thus each image is processed from a sub-aperture. The radar is designed to have an adequate Doppler bandwidth to permit this and if an image is processed using the whole Doppler bandwidth the resulting azimuth resolution is many times finer than the range resolution. Each image processed from a sub-aperture is detected (*ie* the modulus is formed, or an intensity image formed) and then the images are added incoherently. The objective is to produce speckle smoothing, that is, an improvement in the radiometric resolution at the expense of spatial resolution. An explanation of how speckle arises will be found in Ref 24.

Each sub-aperture is generated by means of a digital band pass filter - the 'look filter' - centred on to a different part of the Doppler spectrum. The filtering can be performed before the azimuth correlation process in which case it is known as prefiltering or alternatively, during the azimuth correlation process. In principle it could also be carried out afterwards but this is never done for reasons that will become clear.

Prefiltering has computational advantages. Suppose that the aperture (Doppler spectrum) is divided into M sub-apertures each of width $\Delta F_A/M$ where ΔF_A is the total Doppler spectrum. It is then evident that since we have an

analytic signal it is possible to subsample the Doppler phase and reduce the number of pulses which it is necessary to process. Thus for M sub-apertures it is only necessary to process every Mth pulse.

9.1 Design of the prefilter

The first requirement of the prefilter is that it should not introduce spurious phase shifts into the data, since this will have serious consequences for the azimuth correlation process. A linear phase response (as a function of frequency) might just be tolerable since this would just shift the image in the azimuth direction. It is, however, possible to design a digital filter which introduces no phase shift at all. This is one of the advantages of operating with digital data; a zero phase shift analogue filter is impossible!

There are basically two types of digital filter, the finite impulse response (FIR) and the infinite impulse response (IIR) filters. Only the first type can be designed to have zero or linear phase shifts. It is possible to utilise a filter which has a non zero phase shift by passing the output of the filter back through the filter in reverse time sequence. This removes the phase shifts and is called two pass operation. However, this filter requires all the output pulses to be realised, not just every Mth pulse. In addition this type of filter causes difficulty with the data management. There are similar problems with the recursive filter. An additional problem is that it is not straightforward to change the filter characteristics in 'mid image'.

By far the most straightforward design is the phase-shift free non recursive filter and this is the type which has been used in the RAE processor.

FIR filters are explained in detail in Ref 23.

Consider a sequence $h(n)$ $-\left(\frac{N-1}{2}\right) \leqslant n \leqslant \left(\frac{N-1}{2}\right)$ N odd. These may be regarded as the filter weights. The discrete Fourier transform of this sequence is

$$H(\exp i\omega) = \sum_{n=-(N-1)/2}^{(N-1)/2} h(n) \exp - i\omega n \qquad (103)$$

and this defines a finite Fourier series period 2π and gives the frequency response of a non recursive filter. If there is no phase shift then obviously $H(\exp i\omega)$ is real, and this is true if $h(n) = h(-n)$ and then

$$H(\exp i\omega) = h(0) + 2 \sum_{n=1}^{(N-1)/2} h(n) \cos \omega n \quad . \qquad (104)$$

This argument also works for N even in which case

$$H(\exp i\omega) = 2\sum_{n=1}^{N/2} h(n) \cos \omega n \quad . \tag{105}$$

It is therefore possible to design a filter in which there is no phase shift. It will not be demonstrated here, but the case N even leads to problems because the output time samples occur midway between the input samples. Hence filters used for prefiltering have an odd number of coefficients and have symmetric coefficients.

One way to design the filter is the 'window method'. First of all we decide on the shape of the required response. The ideal (from one point of view) is the rectangular function shown in Fig 18. This is not realistic, however, because it requires an infinit number of Fourier coefficients and hence an infinite sequence of input samples. The resulting Fourier series is truncated therefore and then the coefficients are:

$$h(0) = \frac{\omega_c}{\pi\nu} \tag{106}$$

$$h(n) = \sin \frac{n\omega_c/\nu}{2\pi n} \qquad n \neq 0 \tag{107}$$

where ν = pulse recurrence frequency, ω_c = cutoff frequency (rad/s), and $-(N-1)/2 \leqslant n \leqslant (N-1)/2$.

Due to the Gibb's phenomenon there are ripples in the pass band and in the stop band and the amplitude of these ripples is not affected by the order of the filter N . These ripples may be smoothed by a weighting function and this function works in exactly the same way as the sidelobe weighting function in section 7. Typically, the Hanning weighting function is used

$$w(n) = 1 + \cos \frac{\pi n}{k} \tag{108}$$

where

$$k = \frac{N+1}{2} \tag{109}$$

and then

$$h(0) = \frac{\omega_c}{\pi\nu} \tag{110}$$

$$h(n) = \left[1 + \cos \frac{\pi n}{k}\right] \sin \frac{n\omega_c/\nu}{2\pi n} \quad . \tag{111}$$

Substituting in equation (104) then gives the filter response function:

$$H(\exp i\omega) = \frac{\omega_c}{\pi\nu} + \sum_{n=1}^{N-1} \frac{(1 + \cos n\pi/N) \sin n\omega_c/\nu \cos n\omega/\nu}{n\pi} \quad . \qquad (112)$$

The band centre can be shifted by weighting by an exponential weighting function:

$$h'(n) = h(n) \exp - \frac{i2\pi f_0 n}{\nu} \qquad (113)$$

where f_0 is the shift.

The sub-sampling and weighting procedure is then described by:

$$S'(m) = \sum_{n=-(N-1)/2}^{(N+1)/2} h(n)S(n + Mm) \qquad (114)$$

M is the data reduction factor and S(n), n = 1 ... are complex range compressed pulses. S'(m) are the prefiltered pulses. It is important that the prefilter be as simple as possible since too great a complexity will reduce the computational advantages gained from data reduction. So minimum N consistent with sufficiently low ambiguities is chosen. The ambiguities are caused by non-zero response in the stop band, *ie* aliassing, and they also depend on the data reduction factor. Response curves for this type of filter are given in Ref 11.

Prefiltering has advantages for both frequency domain azimuth processing and for time domain processing. The reduction of the data rate means that corner turning (see section 10) can be done on smaller arrays with consequent time saving and the forward FFT processes a smaller amount of data. The savings for time domain processing can be considerable. The number of multiplications in time domain processing is proportional to the square of the number of pulses, reducing the data rate by 4 (for 4 look processing) reduces the number of multiplications by a factor of 16. Likewise the number of multiplications in performing the FFT is proportional to $N \log_2 N$ and is thus reduced by a factor of 8.

9.2 Frequency domain look filtering

Look filtering can be performed in the frequency domain, and this approach is commonly followed in processors which perform azimuth correlation in the frequency domain, see Ref 9 section 3.3. An azimuth line is assembled for azimuth compression (see section 11 for an explanation of this process) and

Fourier transformed to give the corresponding spectrum. This spectrum then covers the whole Doppler band. If M looks are required then this spectrum could be simply split up into M non zero bands. However, the opportunity can be taken here to introduce sidelobe weighting at the same time. Simply splitting up the spectrum would result in an image azimuth point spread function of 'sin x/x' form assuming there is no weighting applied to either the Doppler spectrum or replica (see later sections). This subject is further explored in the sections on azimuth compression.

10 CORNER TURNING

After range compression each pulse is stored on a direct access bulk storage peripheral - invariably a disc file unit. Such disc file units store the data serially line by line. Moreover the data are usually transferred to and from the disc file in the form of blocks of typically 1024 16-bit words. These blocks are stored at random on the disc wherever there happens to be space. It is evident that if a matrix is stored row by row and it is required to access the columns of the matrix then very many blocks will have to be retrieved from the disc because each block will contain few column elements. Indeed if the length of each row in the matrix exceeds 1024 words (which is invariably the case with SAR data) then one block will have to be transferred for each column element.

In order to set up the data for azimuth compression it is necessary to take samples from each pulse, *ie* to read along the stored 'columns'. Unless the data are organised correctly this process becomes hopelessly inefficient, and consequently time consuming.

There are a number of techniques available for reorganising the data. One of the simplest and more efficient methods is to transpose the data so that the stored blocks contain the data in serial column form rather than serial row form. This process is known as 'corner turning'. This transpose operation can itself be carried out in a number of ways. The technique described by Eklundh[25] is optimum in terms of the number of operations carried out, but its implementation depends on the individual computer system and it has to be modified to handle non square matrices.

Corner turning is not strictly necessary, and indeed it is not used for example in the RAE processor. It is used, however, in both the Jet Propulsion Laboratory processor in the USA[7] and the processor designed by Macdonald, Dettwiler and Associates in Canada[9]. One alternative strategem is to write special low level routines to transfer data on and off disc in large blocks, taking advantage of the structure of the disc file unit to incorporate some parallel reading and writing to and from the columns formed by the tracks on the stack of plattens

which make up a disc file unit. The size of the blocks is limited in various ways by the host computer. When data are read off the disc for azimuth compression, then this is done in large (64 kbyte) blocks. The shape of each block, *ie* the row and column lengths is chosen to maximise the number of new image points which can be processed (in azimuth) using the data taking into account the range migration. Then all the image points which can be processed using this data, and which have not previously been processed, are processed.

The organisation of the azimuth data for both frequency and time domain processors will be discussed further when the effects of range migration are examined.

It should also be pointed out here that if a large enough random access memory is available to store all the range compressed data to process a given image then corner turning is not necessary. At the time of writing such very large memories are just possible but they are expensive. Matters will, however, change and in a few years' time it is probable that corner turning will become a thing of the past. SAR processor designers have this point actively in mind and future designs will undoubtedly feature large fast random access memories.

11 AZIMUTH COMPRESSION WITH NO RANGE MIGRATION

11.1 Frequency domain versus time domain

As pointed out in section 2.2 azimuth correlation is essentially the same process as range correlation. It may be performed either in the time domain or in the frequency domain via the FFT. For a small number of samples correlation via the time domain is faster than via the FFT due to the inherent speed advantage of the FFT method being outweighed by the greater time penalty of memory read/write operations. For larger N, however, the FFT method is much faster, the cross over point depends on the processor, and for the Floating Point Systems AP120B for example, lies between 64 and 128 points. Satellite SAR azimuth processing is performed on data sets of several thousand samples. Hence, when there is no range migration there are clear advantages for processing in the frequency domain. The situation when there is range migration is discussed in section 14.

11.2 Data selection and packing

Let the velocity of the radar platform be V_p and the prf ν. Then the distance between pulses on the ground is V_p/ν metres. If the image has a length L in the azimuth direction then $L\nu/L_p$ pulses are required, ignoring, for the moment, 'end' effects due to: (a) extra pulses at each end of the record being required to ensure that all the pulses which cover the image points at the ends of the azimuth line are included, and (b) to cover end effects in the correlation.

The required resolution determines the Doppler bandwidth ΔF_A - see equation (15). ΔF_A is easily shown to be approximately (*ie* to first order) $2V_p^2 T_A/\lambda a_0$ where T_A is the synthetic aperture time interval, a_0 is the range to a chosen point at $t = 0$, *ie* broadside on in this case. Hence for a constant ΔF_A (*ie* constant resolution) T_A must increase across track with a_0. This happens naturally because the projected width of the antenna footprint increases linearly with a_0 across track, see Fig 19. It does mean, however, that in order to maintain constant azimuth resolution across track the number of samples in the azimuth data (*ie* the number of probes) must be increased as each azimuth line is processed. In addition the number of samples in the azimuth chirp replica $\exp - i2\pi f_0 t_D'$ (equation (13)) must also be increased. Also the contingencies mentioned above to cover the ends of the azimuth line, are also range dependent.

Consider, for the moment, the processing of the whole azimuth spectrum to obtain a maximum resolution one look image. If the azimuth length of the required image is L, the prf ν and the radar velocity V_p then evidently the number of pulses required is $L\nu/V_p$. The number of samples in the Doppler replica against which the signal is correlated is given by the aperture time $T_A = \Delta F_A \lambda a_0/2V_p^2$ and the prf, ν. That is the number of samples = νT_A.

The contingencies should ideally be added to the longer of the two records, *ie* if the number of pulses $n_p > n_r$, the number of replica samples then the contingency (= $n_r/2$ for each of the two contingencies (1) and (2) given above, *ie* n_r), should be added to each end of the n_p record.

In order to perform the correlation via an FFT the number of samples should be 2^n and one has the option of either packing the record length with zeros or choosing a record length (including contingency) equal to 2^n. In fact both the signal record (azimuth line) and azimuth replica expand as each azimuth line is processed moving across track in order to maintain resolution.

If one is prepared to accept a variation in azimuth resolution across track then constant length records are acceptable. In this case it is only necessary to pack one record (probably the replica) whereas in the former case it is necessary to pack both. However variable length records do not constitute a major problem.

11.3 The replica

The replica is $\exp - i2\pi f_0 t_D'$ with t_D' given by equation (13) except that in this section it is assumed that $a_1 = 0$, *ie* there is no range walk. The replica is then the same for all points along a given azimuth line at constant range sample number, the only difference being a shift in the time axis. This

translates into a phase shift in the spectrum and therefore all points contribute the same spectrum (scaled by the scattering cross sections) with different phase shifts. Thus all points may be correlated by the same replica spectrum and, on inverse transforming, these phase shifts translate back into time shifts and each focussed point is returned to its correct sequence in the image line.

The range history is computed for each azimuth line, but is constant for that line once computed. It is emphasised that in this section it is assumed that there is no range migration effect (which is valid for most aircraft SARs for instance). The replica is computed at time increments equal to the time between each pulse and the number of samples corresponds to the number of pulses in the aperture, or sub aperture.

11.4 The correlation process

This is identical to the range compression process described in section 8. The Doppler signal (the azimuth line) is Fourier transformed via an FFT. The range history is computed and Fourier transformed, complex conjugated and the two spectra are multiplied together, multiplied by the sidelobe reduction weights (section 7), inverse transformed, and the image line stored. The end image points in the record corresponding to half the replica length are rejected since they are not properly correlated.

11.5 Multilook processing

So far 'one look' processing with the entire Doppler spectrum has been assumed. In practice multilook images are usually required and each separate look is produced by processing a sub-aperture. This can be done by processing prefiltered data (see section 9) or by selecting a sub-aperture during the correlation process (as is done in Ref 9 for instance).

When processing prefiltered data the processing proceeds exactly as before except that the effective prf is now reduced by the data reduction factor (which equals the number of looks). Thus to process a given number of image points requires fewer pulses. Naturally, each look is processed from a different set of prefiltered data so that the total number of pulses processed remains the same. However overall processing time is shorter because it depends upon the number of pulses in a non-linear way. The effective prf is reduced by the data reduction factor M and there are fewer pulses in the aperture, but since the (sub) aperture is now reduced in width also by a factor M the total number of pulses in the aperture is reduced by M^2 (16 for a 4 look system). Hence if there are approximately 4000 pulses in the full aperture (for SEASAT) there are only 250 effective pulses in one quarter aperture (approximately). For an eight

look system there would only be 63 or so and time domain correlation would begin to become a possibility!

Hence in the construction of the replica the range history is computed at intervals of $\Delta t/M$ where $\Delta t = 1/\nu$ and only for those samples within an interval of T/M where T = full aperture width. Note that each sub-replica is computed using *exactly* the same parameters, the sub-replicas only differ in the range of time over which they are computed. This automatically ensures that the azimuth image lines computed from each sub-aperture are correctly registered. The processing then proceeds as before. The sub-replica is packed with zeros to make it the same length as the azimuth line (in terms of the numbers of samples), Fourier transformed, multiplied by the sidelobe reduction weights, multiplied by the complex conjugate of the Fourier transform of the azimuth line and then inverse transformed. This is repeated for the other sub-apertures. The imperfectly correlated image points at the beginning and end of the image lines are rejected. The modulus of each of the complex valued azimuth lines is taken and they are added incoherently to produce a smoothed image line.

An alternative strategy to prefiltering the data is to generate the looks during the correlation process. Again, the processing proceeds much as in the full aperture case. The pulses covering the required azimuth line (all of them) are selected and a contingency added at each end. They are packed to 2^n and Fourier transformed. The spectrum is then split up into a number of separate contiguous sections. These are the looks. The replicas are constructed exactly as in the case of the prefiltered data and the correlation is then performed in the same way. This technique is used in Ref 9: it will be noted that the sidelobe reduction filter used in Ref 9 was of a different kind from that suggested here, having a linear phase equal ripple Chebyshev response, see section 9.2.

11.6 Azimuth correlation point by point

This technique is very slow in comparison with the block processing inherent in the FFT correlation. It is not limited to a fixed pixel spacing of $V_p/\Delta F_A$.

The pulses covering the target point are selected and the azimuth phase replica computed for the image point for which the data are being correlated. This is a one point correlation and the time for which the replica is computed corresponds to the time coordinate of the imaged point in the image. The replica samples are complex conjugated and sidelobe weighted and the resulted weighted samples are each multiplied by the corresponding azimuth signal samples (of which there are an equal number). The resulting samples are then summed and

this corresponds to the correlated image point. The processor then moves on to the next point (which is arbitrary) and repeats the process.

12 MAPPING AND INTERPOLATION

The previously explained range and azimuth correlation techniques produce an image in slant range at fixed pixel spacings of $c/2\Delta F_R$ and in azimuth at a fixed pixel spacing of $V_p/\Delta F_A$ where c and V_p are the velocities of light and the radar platform, and ΔF_R, ΔF_A are the video (chirp) bandwidth and Doppler (sub-aperture) bandwidth. See equation (15) and subsequent comments.

There are two difficulties here. First, image users invariably require the image to be an orthographic map projection and not mapped into slant range. Second, arbitrary pixel spacing is desirable not the least because the image detection process (*ie* finding the modulus) increases the effective bandwidth when the data is converted from complex form to real form. It is therefore necessary, in principle, to increase the image sampling rate.

The explicit way to do this is simply to interpolate between image sample points using, for example, cubic splines (see Ref 26 p.349) or by convolving with an interpolating function which, in the case of a SAR image, would be the point spread function. In fact it is far better to interpolate an image *before* detection, *ie* by interpolating the complex valued image. At least, then, the image is correctly sampled before the interpolation begins! There are several possibilities and one of the most practical is to weight the point spread function of the system by several samples on either side of the point at which the image is required to be interpolated, add the point spread functions and evaluate. This is just a limited convolution over a few samples.

One technique which is in use at RAE in the range direction is that explained in section 5.5. After range compression an inverse FFT is performed of k times the initial length. This automatically generates $k - 1$ intermediate samples between the original samples. The subsequent data selection then chooses range samples nearest to the desired ones. This technique is a very fast way of performing interpolation but it requires a lot more memory. It is a powerful technique when used in conjunction with data which has range walk because these intermediate samples reduce ghosting (see section 15) and in addition the range walk implicitly interpolates the data (see section 13).

Most interpolation methods give trouble at the ends of a data set. In principle given an infinite data set sampled at equal intervals the original bandlimited fraction can be completely reconstructed from its samples (see Ref 16, p 141), simply by convolving the data with a 'sin x/x'

function. In practice one has a finite data set. It can be shown that convolving such a finite data set with a 'sin x/x' function is equivalent to fitting a polynomial of degree $n-1$ to the n samples. (Actually this is Lagrangian interpolation, see Ref 26 p.235). The difficulty is that such polynomials 'wiggle' and although the interpolation may be smooth and accurate in the middle of the data set, at the edges it could be a long way out. Fortunately the ends of the data set are not properly correlated anyway due to end effects on the circular convolution and in practice, bearing in mind that the end points of the correlated data are discarded anyway, this method seems to work very well. It will be recognised that inverse Fourier transforming with a larger interval than the forward transform effectively convolves with a 'sin x/x' function and evaluates in between the original samples. The situation is also helped by sidelobe weighting applied to range compression which keeps down the sidelobes in the point spread function and hence the wiggles in the interpolating polynomial.

In the azimuth direction one can also interpolate implicitly by performing the azimuth correlation point by point.

13 RANGE MIGRATION

The Doppler effect is a result of the changing phase between the radar and the target as the radar moves past. The phase $\phi(t)$ is given by

$$\phi(t) = \frac{4\pi}{\lambda}\left(a_0 + a_1 t + a_2 t^2 + \ldots\right) \qquad (115)$$

a_0, a_1 and a_2 are defined after equation (10).

The changing phase is a result of the change in slant range between radar and target. If the change in range is smaller than approximately one half of the range resolution length over the whole synthetic aperture then the samples used in the azimuth compression process (the azimuth line) can be taken from the same sample number (or range gate) from each pulse. The factor of one half is, introduced by the double path length.

On the other hand if the change in slant range is greater than one half resolution cell then the azimuth samples used in azimuth compression have to be taken from different range gates in each pulse. This effect is known as range migration and has already been discussed from a different point of view in section 2.3. Evidently the locus of a point through the range gates (sample space) is very nearly parabolic since the Doppler shift is very nearly a linear function of time with a constant offset. The shape of the parabola is given by equation (115) and contains a linear component (the a_1 term) and a quadratic component (the a_2 term). In this Report the linear component is called range

walk and the quadratic component (the a_2 term) is called range curvature. Together they constitute range migration.

The behaviour of the a_1 and a_2 components were investigated in sections 4.3 and 4.4. Equation (66) shows that the range walk component is caused by orbit eccentricity, squint and Earth rotation. The rotation term is dependent on the latitude because the Earth's surface velocity is greatest at the equator and smallest at the 'top' of the orbit (nearest the poles) because of the much smaller projected Earth radius. In addition the included angle between the swath velocity vector and the Earth's surface velocity varies, being equal to the orbit inclination angle at the equator (if the squint is zero) and zero nearest the poles. The range walk is hence a function of latitude and is not constant in the along track direction, although it can be considered so for a small enough interval. The range walk is also a function of a_0 and of θ_0, *ie* across track, since the grazing angle varies and hence the resolved component of velocity.

The quadratic component is also a function of latitude although a much weaker function than the linear component.

Hence in the case of an orbital SAR the range migration consists of two components (1) a linear component (range walk), and (2) a quadratic component (range curvature). The first is a result of the rotating Earth, non circular orbit and squint and the second is primarily a result of the changing geometry. Note that curvature could also be caused by higher order terms a_3, a_4, etc. These have been ignored here; the cubic curvature term a_3 was investigated in section 4.4 and for the purposes of range migration correction it is negligible.

The situation is made clearer by Fig 20 (which has been adapted from Ref 8 Fig 1). In Fig 8 Δw is the total range walk and Δc is the range curvature.

The range to a target, is given by:

$$a(t) = a_0 + a_1 t + a_2 t^2 \quad . \tag{116}$$

For the example discussed in sections 4.3 and 4.4 a_1 = 65.66 m/s and a_2 = 28.09 m/s^2. For SEASAT the distance between range gates (samples) is about 7.89 m. Hence over an interval of 2 s, corresponding to the maximum synthetic aperture, the total linear range walk Δw is about 65.66 × 2 × 2/7.89 range samples = 33.3. The extra factor of 2 is, again, due to the double path length. In the case of the curvature we take half of the interval or 1 s and this gives Δc = 7.1 range gates. Note that the range walk can be much greater nearer the equator.

The range migration problem is the dominant problem in the design of digital SAR processors for satellite data; and if it can be ignored it very greatly simplifies the processor design. For the future it may be that new technology will permit different satellite systems in which the effect can be ignored, if so it will be possible to much simplify the processor design.

14 AZIMUTH CORRELATION WITH RANGE MIGRATION

Azimuth correlation with range migration can either be performed in the frequency domain or in the time domain. The frequency domain approach is examined first as this is the most usual technique. The method described in this section is that used in the JPL processor, see Refs 7 and 8, it is also employed in the MDA SEASAT processor, see Ref 9.

14.1 Azimuth correlation in the frequency domain

The range migration curve shown in Fig 20 is shaped like a banana and is therefore commonly known at RAE as the data 'banana'. That term, being short and succinct, is used here. The box of data containing the banana, that is the rectangular box of data within which the banana can be fitted is called the 'banana box'.

The first essential idea behind the simultaneous processing of a number of azimuth image points via the FFT is that the banana only slowly varies along the azimuth direction. That is a number of consecutive targets all having the same a_0 have nearly the same banana. Obviously the bananas are shifted in azimuth relative to each other but the shape of each banana is approximately the same. This is shown in Fig 21.

The second essential idea is that there is a one to one correspondence between the azimuth coordinate (in time) and the corresponding Doppler frequency of a target. This is obvious from equations (8) and (9). Therefore if a number of bananas are Fourier transformed, the same points in each banana map onto the same point in Doppler frequency/range space. This is shown in Fig 22. One can then define a spectrum banana and a corresponding spectrum banana box.

Hence if one has a box of data containing many bananas corresponding to many targets all with the same a_0, and these bananas are Fourier transformed the spectra map on top of each other with relative phases corresponding to their relative displacements. The spectrum banana is then formed. Evidently other data bananas corresponding to different values of a_0 are Fourier transformed in the same way and mapped one to one into the spectrum box.

In practice one does not wish to Fourier transform along a banana but across an azimuth line at constant range. If this is done it will be seen from the above arguments that the spectrum bananas are automatically created in the

spectrum/range space due to the one to one correspondence between offset along the banana, range and Doppler frequency.

It is straightforward to identify a spectrum banana corresponding to a group of data bananas because the spectrum banana is simply a data banana stretched out in the azimuth direction to fit the spectrum width.

This elegant technique will only work if all the data bananas have the same shape and herein lies a difficulty. This requirement sets the maximum azimuth distance which can be correlated simultaneously. The data are thus divided up into sections. Each section has a set of different bananas with different range walk gradient. The sections must overlap in order to ensure that all target points are properly correlated (end effects).

Once the box of data has been selected and Fourier transformed line by line the spectrum bananas are obtained by piecewise approximation. A section of spectrum is taken from the appropriate part of each line. Each section of spectrum is then coherently assembled to form a complete spectrum banana - see Fig 21. The appropriate section of spectrum from each line is selected by taking that part of each line which has within one half range gate of the required spectrum banana. Note that all of this is possible because of the one to one correspondences noted above and the linear nature of the Fourier transform.

After the spectrum has been assembled it is operated on in exactly the same way as in section 11, to which the reader is now referred. Briefly, the 'looks' are generated, unless the data is prefiltered, and the replica spectrum is generated via the range history and Fourier transformed, conjugated the spectra multiplied and inverse transformed. The resulting image line is, of course, mapped on to the image along a *straight* line at the corresponding value of a_0 .

14.2 Pre-skewing

It is evident that the number of spectrum segments required is largely dependent on the range walk. An aperture may have a walk of several tens of range gates and there will then be several tens of spectrum segments. It is desirable to decrease the number of segments to improve the speed of the processing. One way to do this is to pre-skew the data by rearranging the range gates relative to each other - see Fig 24. The rate of displacement of the range gates is chosen to be equal to the range walk. It is then only necessary to produce a segmented approximation to the range curvature instead of both the walk and the curvature. This greatly reduces the number of segments.

This technique, however, has a severe disadvantage. The range walk is latitude dependent as pointed out in section 13 and the data must therefore be

divided into sections as in section 14.1. Unfortunately there are difficulties in fitting the sections together with the pre-skewing method because of the different range walk gradient from section to section. There are small gaps between sections or alternatively non registering overlaps and extra manipulation is required to make the sections fit together.

This problem does not arise if the range migration correction is done entirely in the frequency domain - see section 14.1.

14.3 Azimuth correlation in the time domain

This is the technique used in the RAE processor[10]. It is the most straightforward way to correlate data which has severe range migration. Each image point is correlated separately. The azimuth data is selected for the image point of interest by computing the data banana and then selecting each range sample lying within a specified fraction of a range gate of the required banana. If the range compressed data is not interpolated this fraction is one half, see Fig 25. Thus in contrast to the frequency domain correlation, it is the data banana which is piecewise approximated. When the data banana has been assembled the range history of the point is computed at the time corresponding to the relevant time coordinate in the image. The banana samples and the complex conjugate of the replica samples are then multiplied together and this gives the single point correlation for this image point. The process is repeated for the next image point with a new banana and a new replica. The next image point can be spaced completely arbitrarily from the preceding point both in azimuth and range. The range migration effect helps to interpolate the data implicitly.

It will be seen that this technique is very precise and involves no fundamental approximation. There is the piecewise approximation of the data banana but the effects of this can be mitigated by interpolating the range samples (using the Fourier transform interpolating technique described in section 5.5).

Time domain correlation is slow. However, there are considerable advantages. First, the sample spacing of the image is completely arbitrary. One result of this is that the image can be mapped in any way without interpolation. Second, the image can be rotated, the x and y axes of the image do not necessarily have to align with the range and azimuth directions. Third, there are no awkward problems involved in fitting sections together. Above all the image is precisely correlated. All these advantages together produce images of very high quality.

15 GHOST IMAGES

Ghosts are faint replicas of the main image displaced in the azimuth direction and appear on either side of the main image. Fig 26 shows the image of a point target (it is the same image as Fig 1, but without ghost suppression) and its attendant ghosts. There are usually many ghosts visible and they are spaced out from the main image at regular intervals. The position of the ghosts depends upon the sub-aperture position and Fig 26 shows three images from three adjacent sub-apertures added. The ghosts then appear in sets of three. Fig 26 only shows clearly the fundamental ghosts, but in fact there are other groups of three spaced out further from the central point which are not easily visible. These ghost artifacts are produced during the azimuth correlation process and are caused by the piecewise approximation of the spectrum banana in frequency domain processing or the piecewise approximation of the data banana in the time domain, see section 14 for an explanation of these approximations. Ghosts are produced by both frequency domain and time domain processors.

The artifact is briefly analysed in the section by considering the image of a single point on an absorbing background. Fig 27 shows the effect of the piecewise approximation in the time domain, and Fig 28 shows the corresponding effect on the frequency domain. In both cases the data cycles over the top of the range point spread function and hence the data assembled for azimuth compression is amplitude modulated. There is also a phase modulation, but this will be ignored here. The peak to peak ripple amplitude for a 'sin x/x' point spread function in the range direction and range sample spacing equal to the distance between the zeroes of the point spread function ($= c/2\Delta F_R$) is 0.363 of the amplitude of the psf, *ie* -8.8 dB. The fundamental frequency of the modulation is given by the rate at which the data cycles over the point spread function. The analysis which follows is applicable to a time domain processor, it can, however, be easily modified to make it apply to the frequency domain. The range is given by

$$a(t) = a_0 + a_1 t + a_2 t^2 \quad . \tag{117}$$

The distance between range samples is $c/2\Delta F_R$ and so the frequency is given by

$$\frac{2\Delta F_R}{c}\left[a_0 + a_1 t + a_2 t^2\right] \quad . \tag{118}$$

The amplitude modulation produces coherent sidebands in the time domain data.

The Doppler information is contained in the sidebands and these are also correlated but being shifted in frequency from the main Doppler band they are shifted in the azimuth direction in the image. These coherent sidebands are, then, the source of the ghosts in the image. Unfortunately in order to give a satisfactory analysis of the ghosts it is necessary to include second-order effects to explain the observed shift of the ghost positions with sub-aperture position. Such an analysis can be performed but is too long to present here (see Ref 10 for additional information).

The most elegant way to reduce the effect of ghosting is to interpolate the range compressed data. This reduces the period of cycling on the piecewise approximation, *ie* increases the frequency. As a result the cycling occurs over a smaller segment of the point spread function and this greatly reduces the amplitude of the ripple. The type of interpolation referred to here is simply the generation of intermediate samples between the original samples; this is much faster than other types of interpolation and in the RAE processor the intermediate samples are generated via the FFT technique described in section 5.5 With three intermediate points between the original samples the ghosts are reduced to a level of about -35 dB, see Fig 1.

16 CONCLUSION

In this Report an attempt has been made to set down the theory of digital SAR processing for orbital SAR. However, the reader should note that many topics have not been included. For example there are several other azimuth compression techniques. In this connection there is a technique which at RAE is called 'coherent subimaging'. This is a hybrid time domain-frequency domain technique in which a large number of prefilters, say 16 or 32, select Doppler bands and an image is generated from each set of data. The resulting images are *not* detected but are added coherently to produce an image which has full resolution. The technique has several variations, the point being that because the data is prefiltered it can be subsampled and all the benefits of data reduction obtained. Because there are now far fewer pulses in each sub-aperture the time domain approach becomes efficient - more efficient than frequency domain processing - see the comments in section 11 on multilock processing.

Autofocussing techniques have not been addressed, nor have the effects of partial coherence. Another problem which has not been addressed is the problem of dynamic range in SAR processing, that is, how many bits does one need at various stages of the processing to represent the data?

Many of these topics are under active investigation at RAE and this list will serve to demonstrate that the study of methods for digital SAR processing is a problem that is certainly not completely solved; much work is still needed in many areas, for example on achieving higher speed processing, on selection of optimum radar parameters and on special purpose processing to enhance particular image features.

REFERENCES

No.	Author	Title, etc
1	R.O. Harger	Synthetic aperture radar systems. Academic Press (1970)
2	S.A. Hovanessian	Introduction to synethetic array and imaging radars. Artech ISBN 0-89006-082-7 (1980)
3	J.J. Kovaly	Synthetic aperture radar. Artech ISBN 0-89006-056-8 (1976)
4	W.M. Brown G.G. Houser R.E. Jenkins	Synthetic aperture processing with limited storage and presumming. IEEE Trans Aerospace and Electronic Systems AES-9 No.2, pp 166-176 (1973)
5	J.C. Kirk	A discussion of digital processing in synthetic aperture radar. IEEE Trans Aerospace and Electronic Systems AES-11, No.3, pp 326-337 (1975)
6	W.J. Van de Lindt	Digital technique for generating synthetic aperture radar images. IBM J. Res. Develop., Vol 21, No.5, pp 415-432 (1977)
7	C. Wu	A digital system to produce imagery from SAR data. AIAA Paper 76-968. In Proc AIAA Systems design driven by sensors conference. Pasadena, Ca (1976)
8	C. Wu B. Barkan W.J. Karplus D. Caswell	Seasat synthetic aperture radar data reduction using parallel programmable array processors. IEEE Trans Geoscience and Remote Sensing, Vol GE-20, No.3 (1982)
9	J.R. Bennett I.G. Cumming	A digital processor for the production of Seasat synthetic aperture radar imagery. SURGE Workshop, ESRIN Frascoti, Italy, 16-18 July 1979
10	D.G. Corr	Experimental SAR processor study and system implementation manual. Systems Designers Ltd Report C0875 prepared under Contract A926/502 for Space and New Concepts Dept, RAE, Farnborough

REFERENCES (continued)

No.	Author	Title, etc
11	S.R. Brooks A.P. Luscombe A.B.E. Ellis	Marconi Research Co Final Report on-board SAR processor definition. Report MTR 79/96 prepared under Contract 2897/76 for ESTEC. ESA report reference ESA CR(P) 1284
12	J.R. Bennett I.G. Cumming J. Lim R.M. Wedding	A fast, programmable hardware architecture for the processing of spaceborne SAR data. 17th International Symposium on Remote Sensing of the Environment Arm Arbor, Michigan, 9-13 May 1983
13	K. Tomiyasu	Tutorial review of SAR with applications to the imaging of the ocean surface. Proc IEEE Vol 66, pp 563-583 (1978)
14	W.M. Brown L.J. Porcello	An introduction to synthetic aperture radar. IEEE Spectrum, pp 52-62 (1969)
15	R.K. Raney	Processing synthetic aperture radar data. Int. Jnl. Remote Sensing, Vol 3, No.3, pp 243-257 (1982)
16	A. Papoulis	Signal analysis. McGraw-Hill ISBN 0-07-048460-00 (1977)
17	J.L. Synge	Relativity: The special theory. North-Holland (1956)
18	D. King-Hele	Theory of satellite orbits in an atmosphere. Butterworths (1964)
19	E. Cutting G.H. Born J.C. Frautnick	Orbit analysis for Seasat-A. Jnl. of the Astronautical Sciences, Vol 26, No.4, pp 315-342 (1978)
20	E.C. Titchmarsh	Introduction to the theory of Fourier integrals. OUP (1937)
21	J.R. Klauder A.C. Price S. Darlington W.J. Albersheim	The theory and design of chirp radars. Bell System Technical Journal Vol 39, No.4, pp 745-808 (1960)

REFERENCES (concluded)

No.	Author	Title, etc
22	G.D. Bergland	A guided tour of the fast Fourier transform. IEEE Spectrum, pp 41-52 (1969)
23	L.R. Rabiner B. Gold	Theory and application of digital signal processing. Prentice-Hall ISBN 0-13-914101-4 (1975)
24	B.C. Barber	Some properties of SAR speckle. Chapter 8 in Satellite Microwave Remote Sensing Editor T.D. Allan, Ellis Horwood, ISBN 0-85312-494-9 (1983)
25	J.O. Eklundh	A fast computer method for corner turning. IEEE Trans on computers, Vol C-21, pp 801-803 (1972)
26	R.W. Hamming	Numerical methods for scientists and engineers. McGraw-Hill, 2nd Ed, ISBN 0-07-025887-2 (1973)

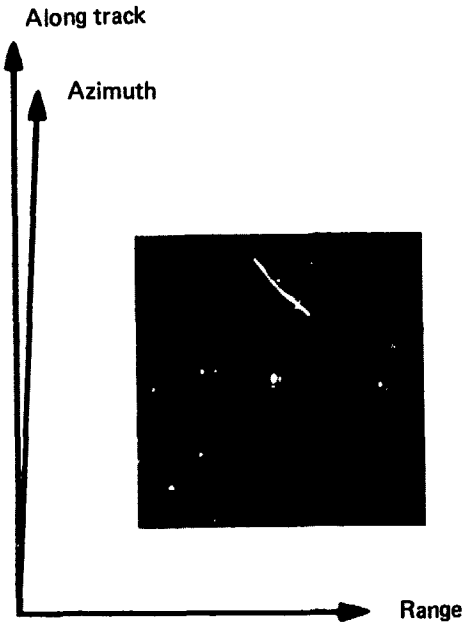

Fig 1 Image of point target (Goldstone corner reflector from SEASAT rev.882) showing the difference between along track and azimuth directions

Fig 2

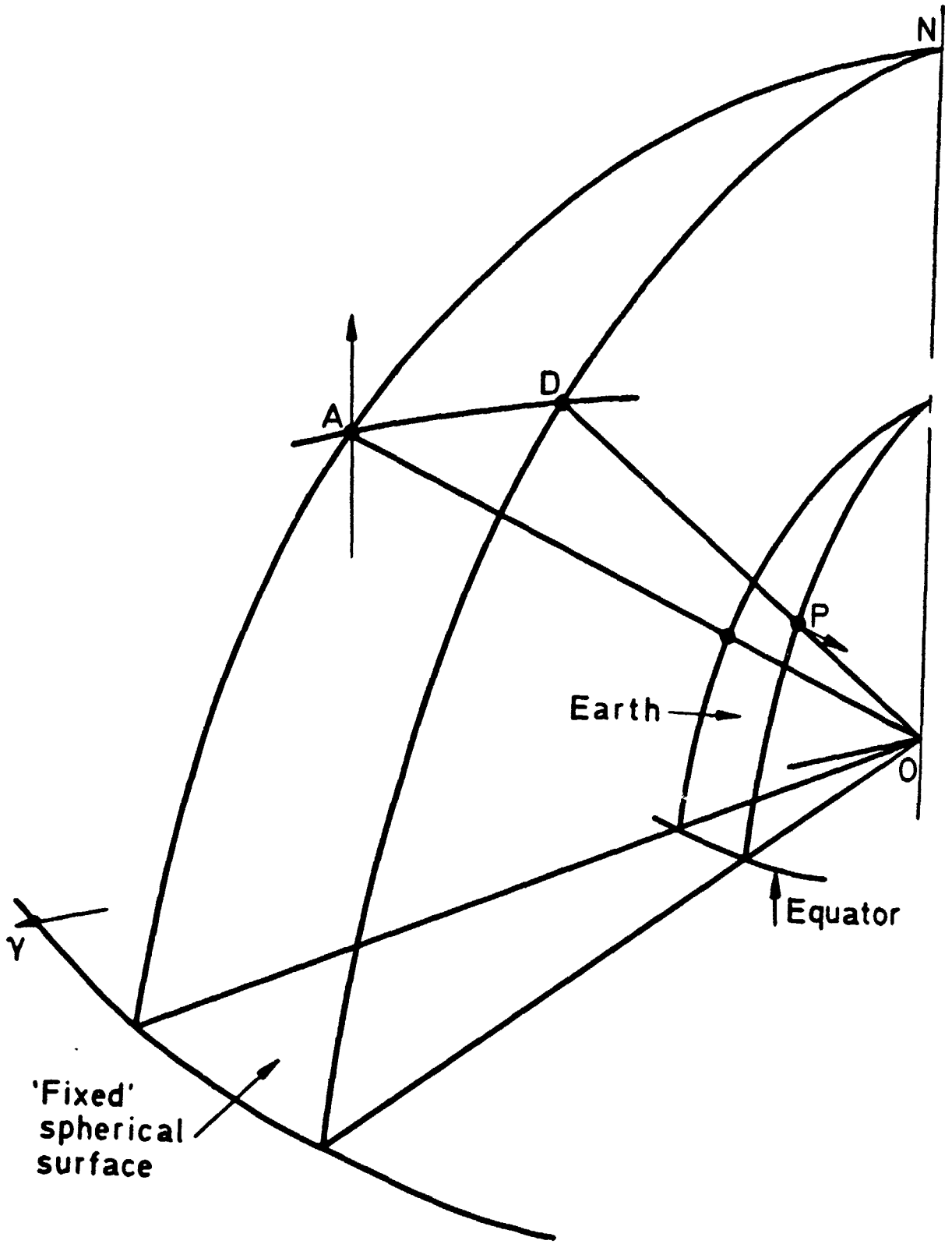

Fig 2 Projection of typical orbit onto a fixed sphere

TR 83079

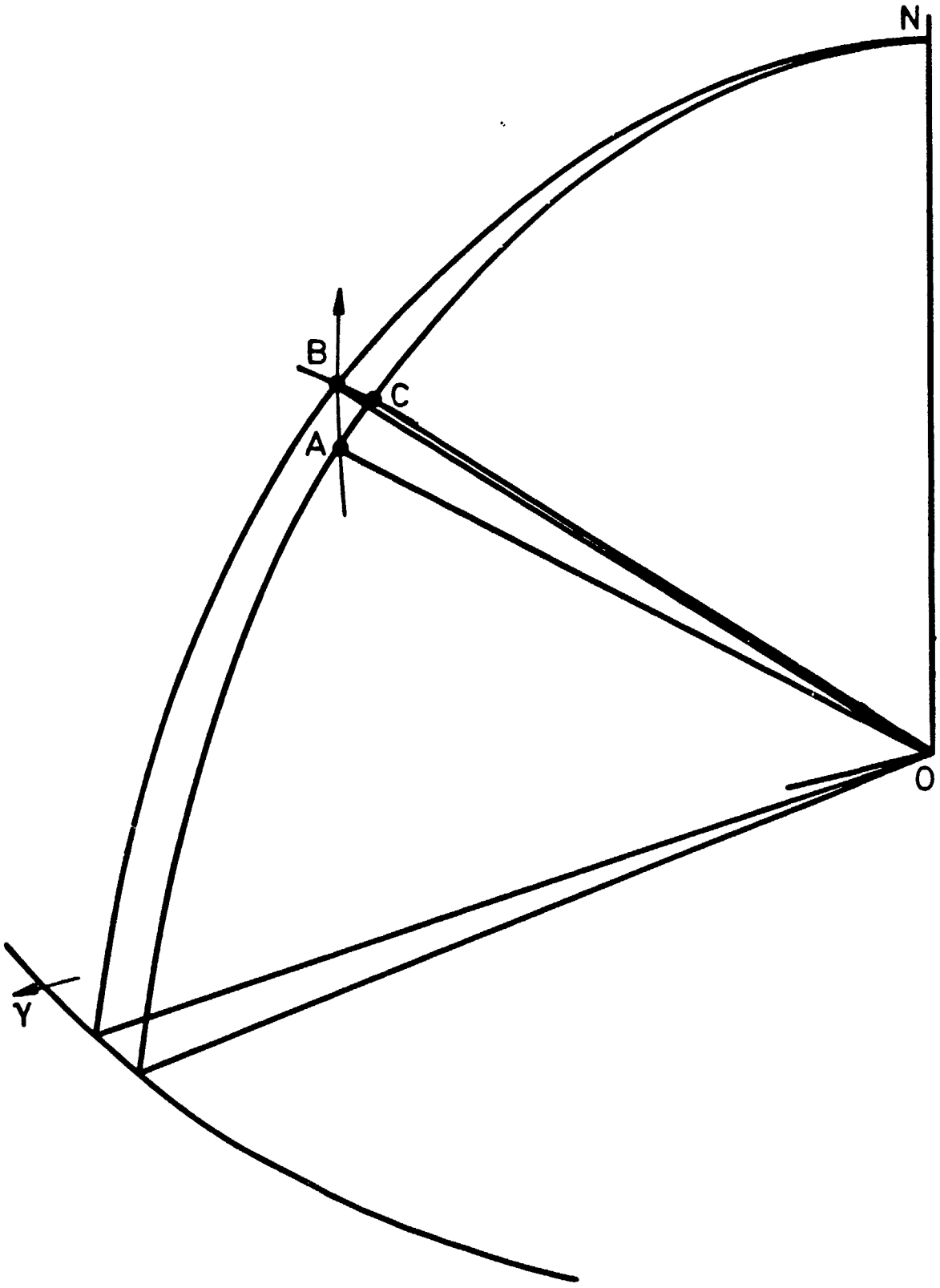

Fig 3 Change in position of the satellite's projection

Fig 4 SEASAT image of the Wash and East Anglia from rev.762

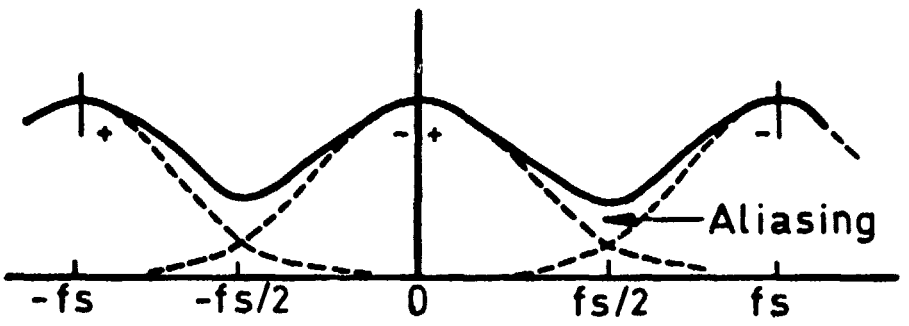

Fig 5 Periodically repeated spectrum with aliasing

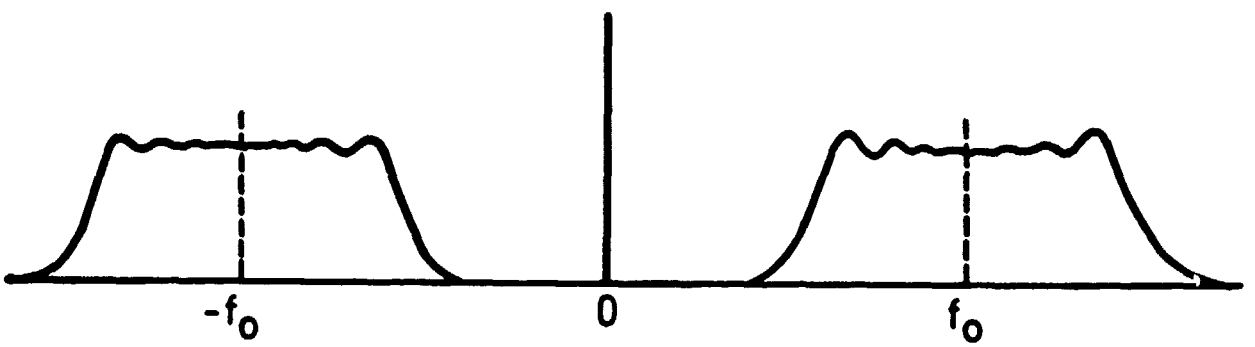

Fig 6 Spectrum of real valued chirp centred on F_0

Fig 7 Spectrum of received real valued signal centred on F_0

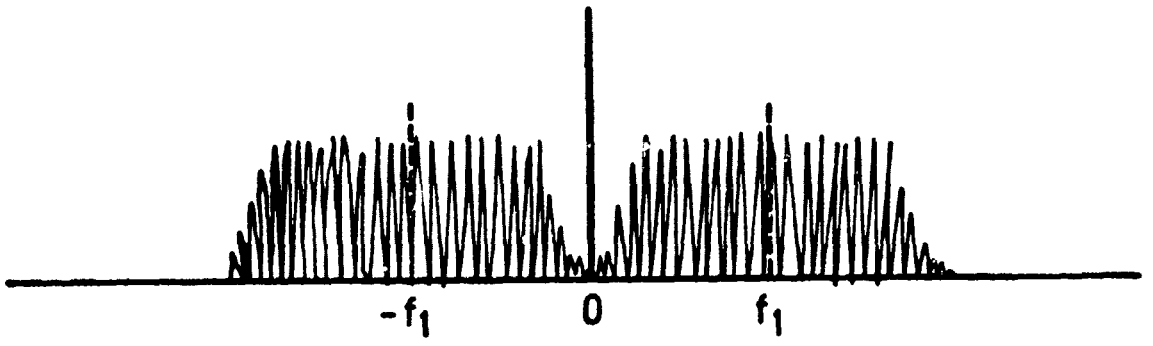

Fig 8 Spectrum of received real valued signal centred on F_1

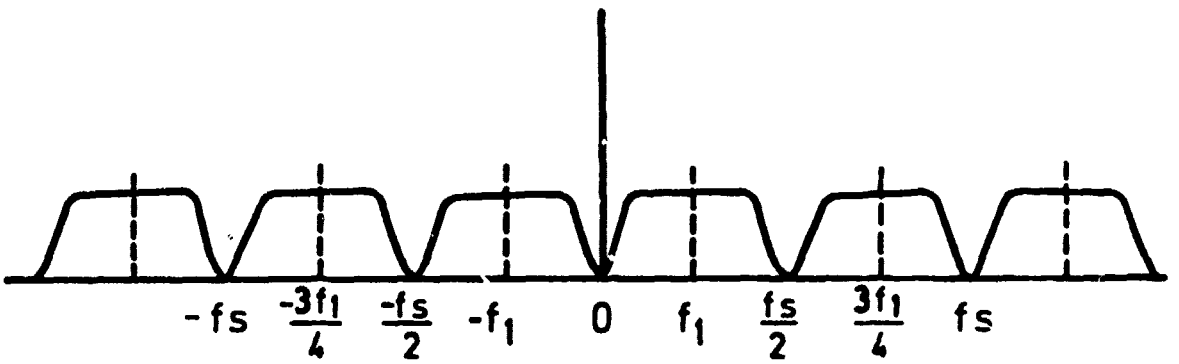

Fig 9 Periodically repeated range spectrum

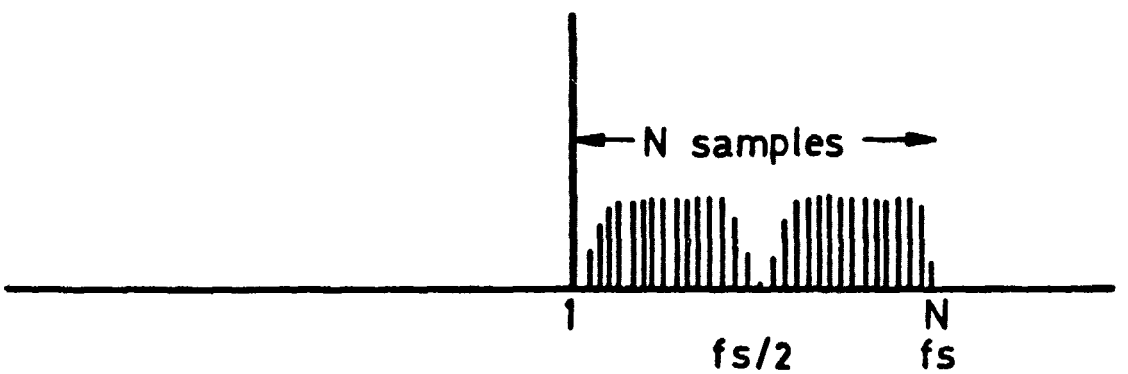

Fig 10 Sampled spectrum for range compression

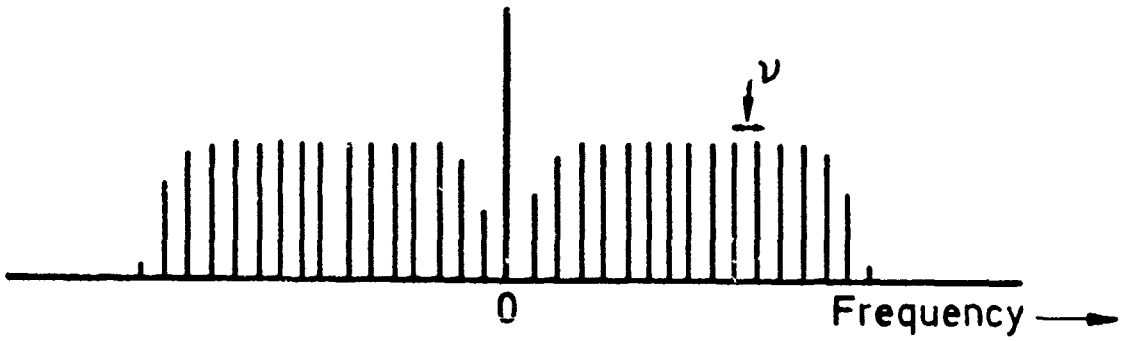

Fig 11 Line range spectrum caused by periodic pulsing of the radar

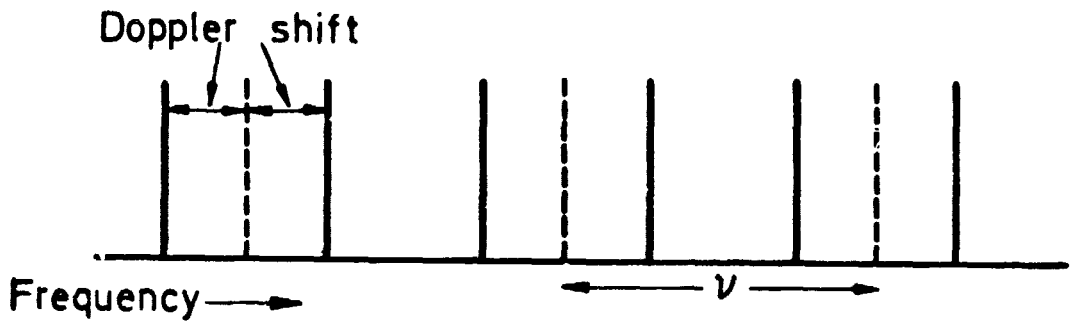

Fig 12 Line split due to Doppler shift with cosine phase term

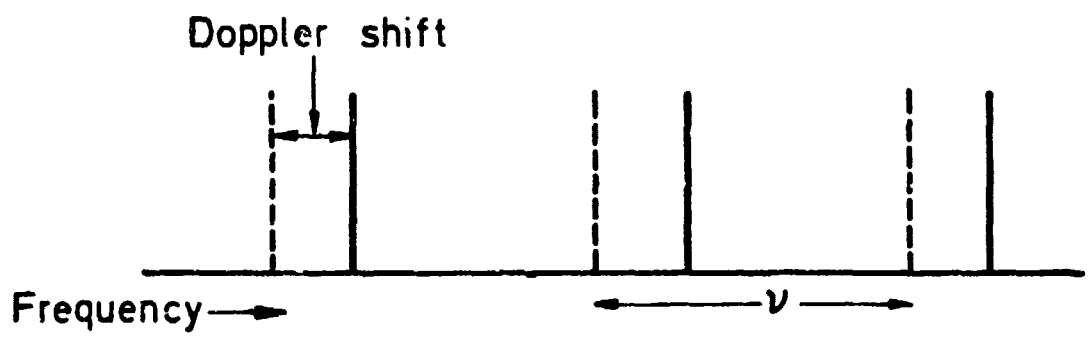

Fig 13 Line shift due to Doppler shift with exponential phase term

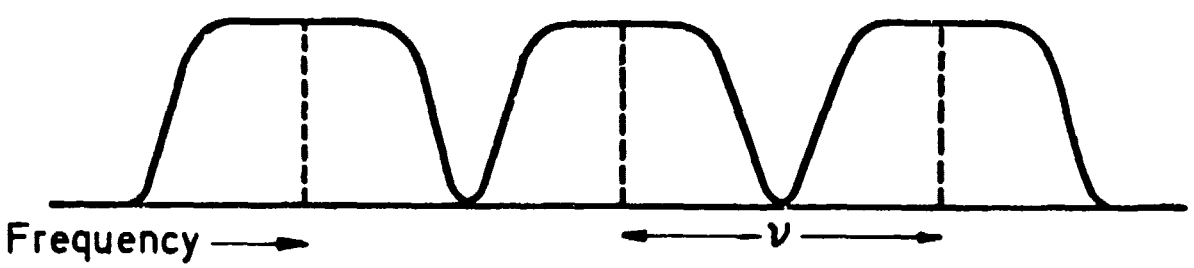

Fig 14 Spectral lines broadened into a Doppler continuum

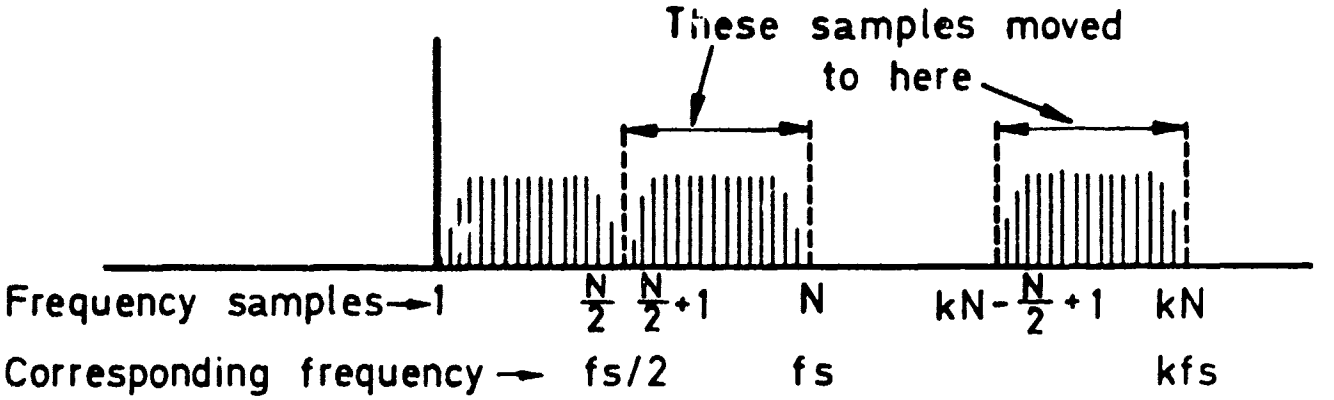

Fig 15 Spectrum of interpolated data

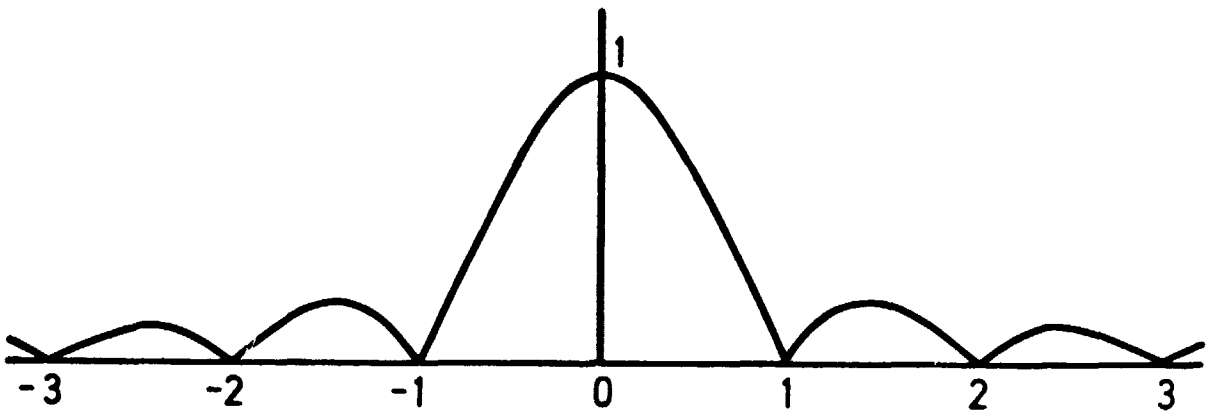

Fig 16 Modulus of point spread function $|\sin \pi x/\pi x|$

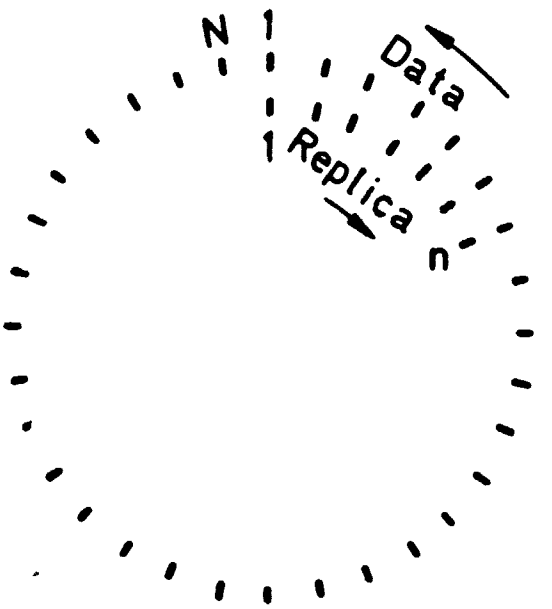

Fig 17 Cyclic correlation

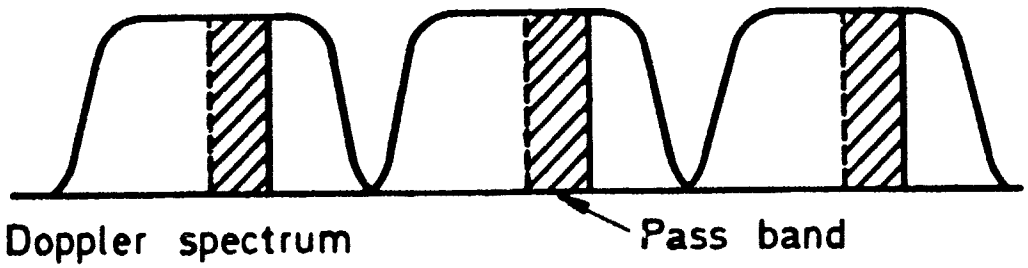

Fig 18 Rectangular pre-filter response

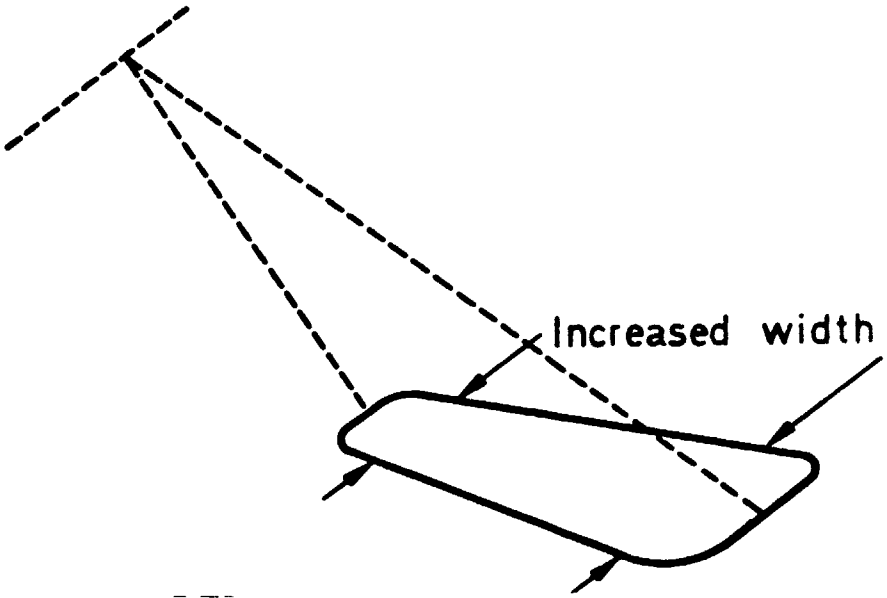

Fig 19 Projected width of antenna footprint

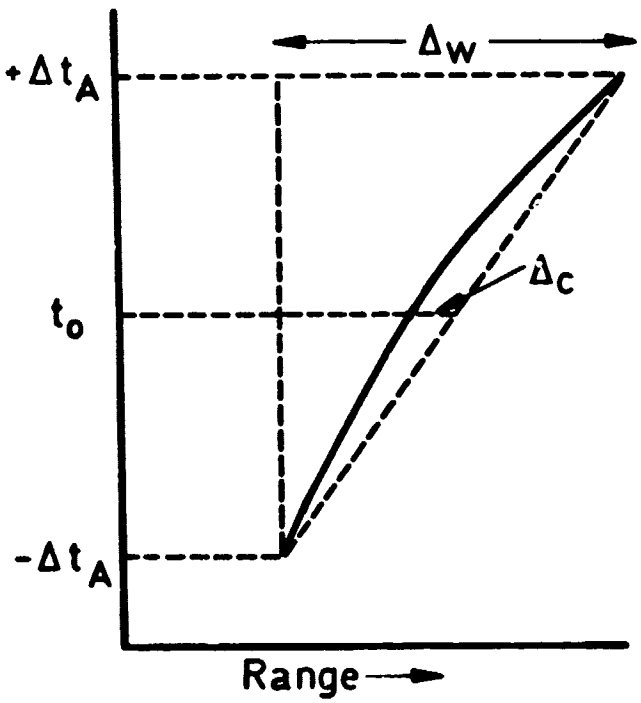

Fig 20 Range walk and curvature

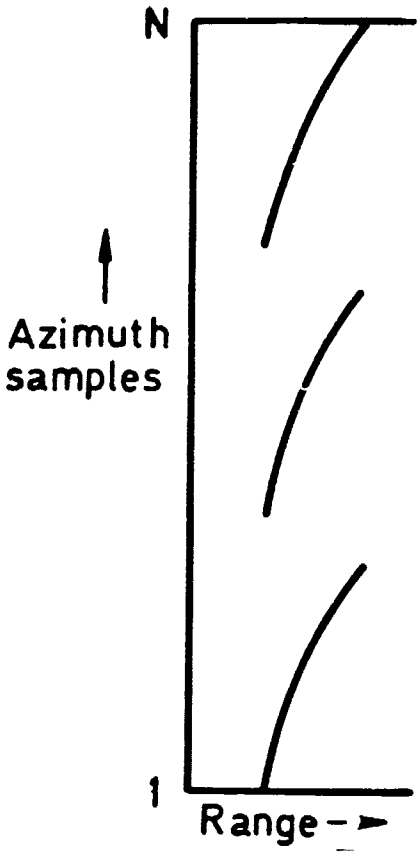

Fig 21 Data bananas

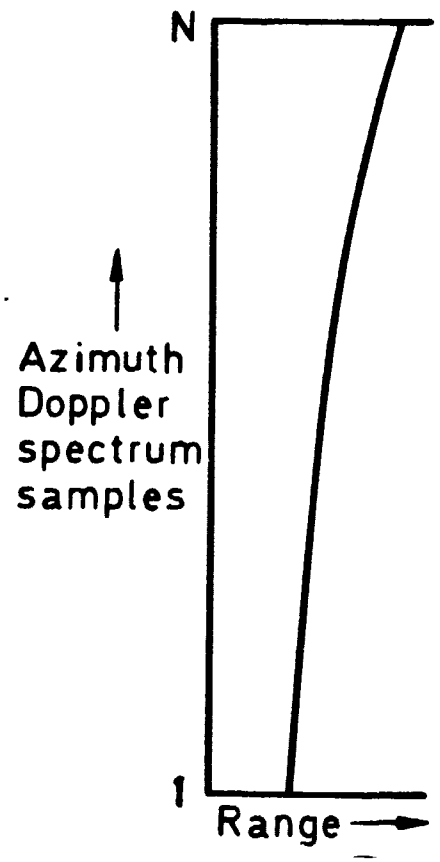

Fig 22 Corresponding spectrum bananas

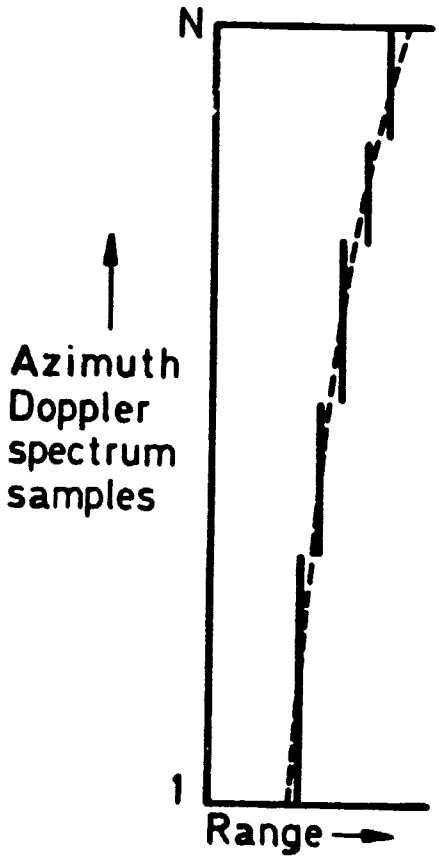

Fig 23 Piecewise approximation of the spectrum banana

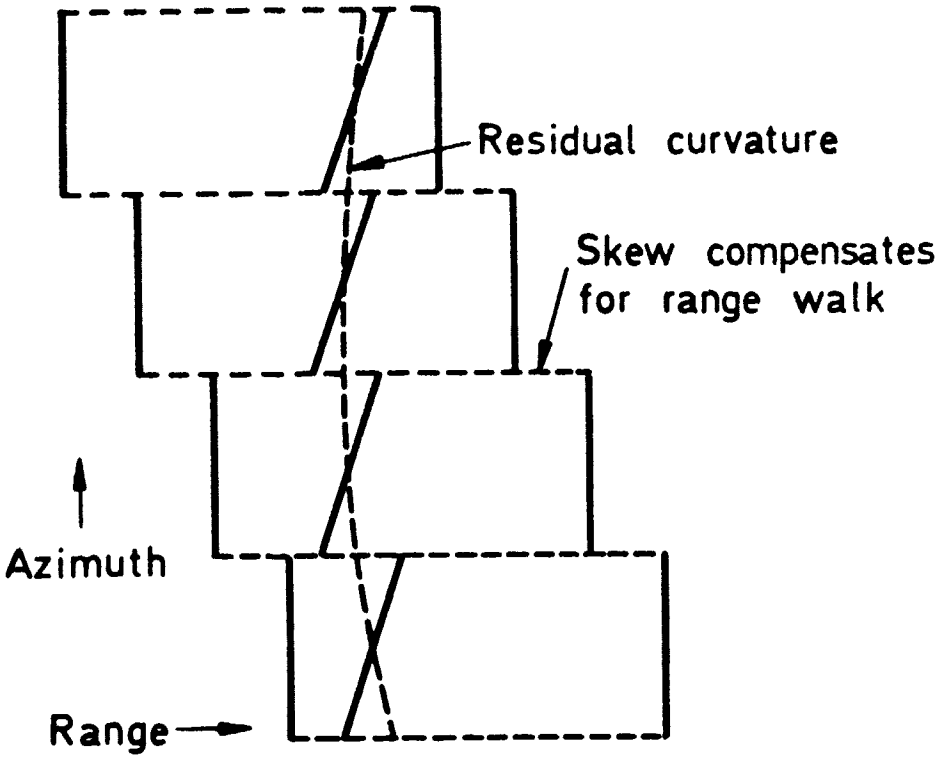

Fig 24 Skewing the data to compensate range walk

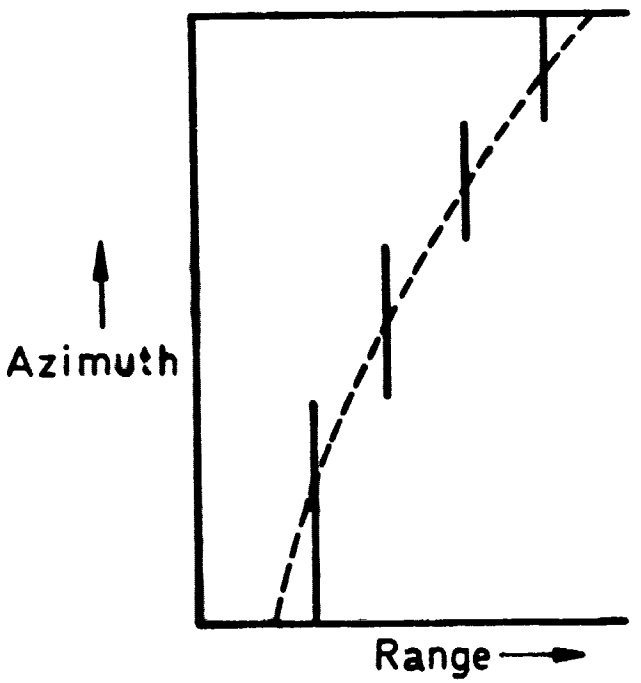

Fig 25 Piecewise approximation of the data banana

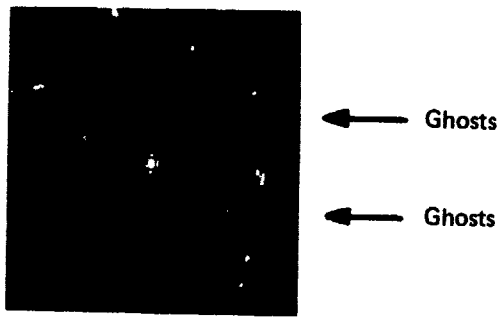

Fig 26 Image as in Fig 1 but with no ghost suppression

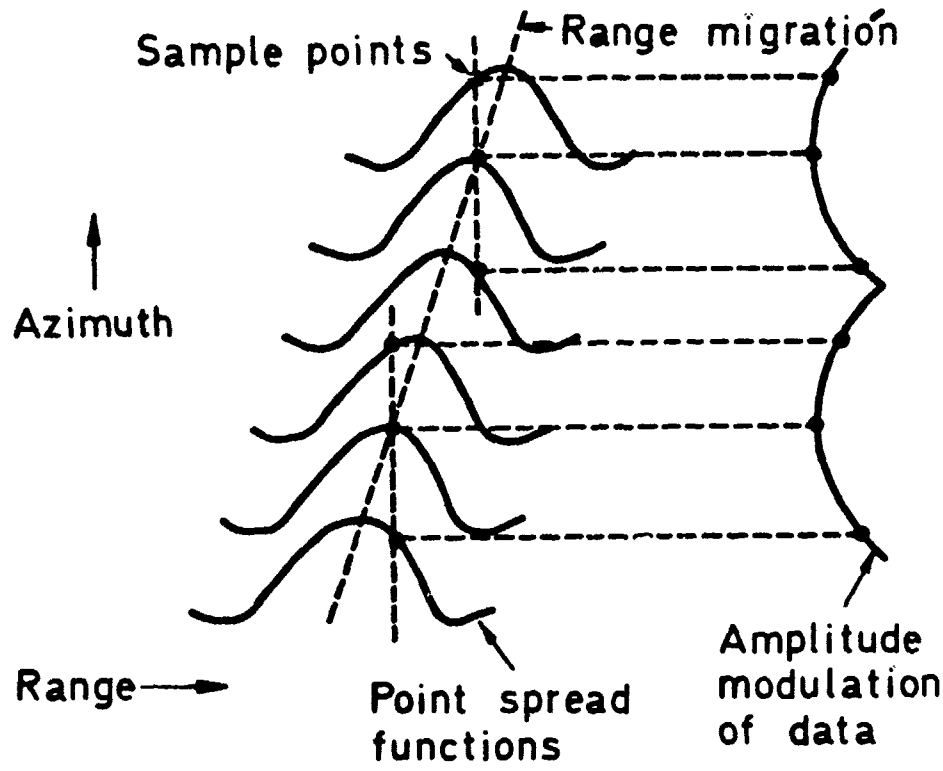

Fig 27 Amplitude modulation of the data due to piecewise approximation in the time domain

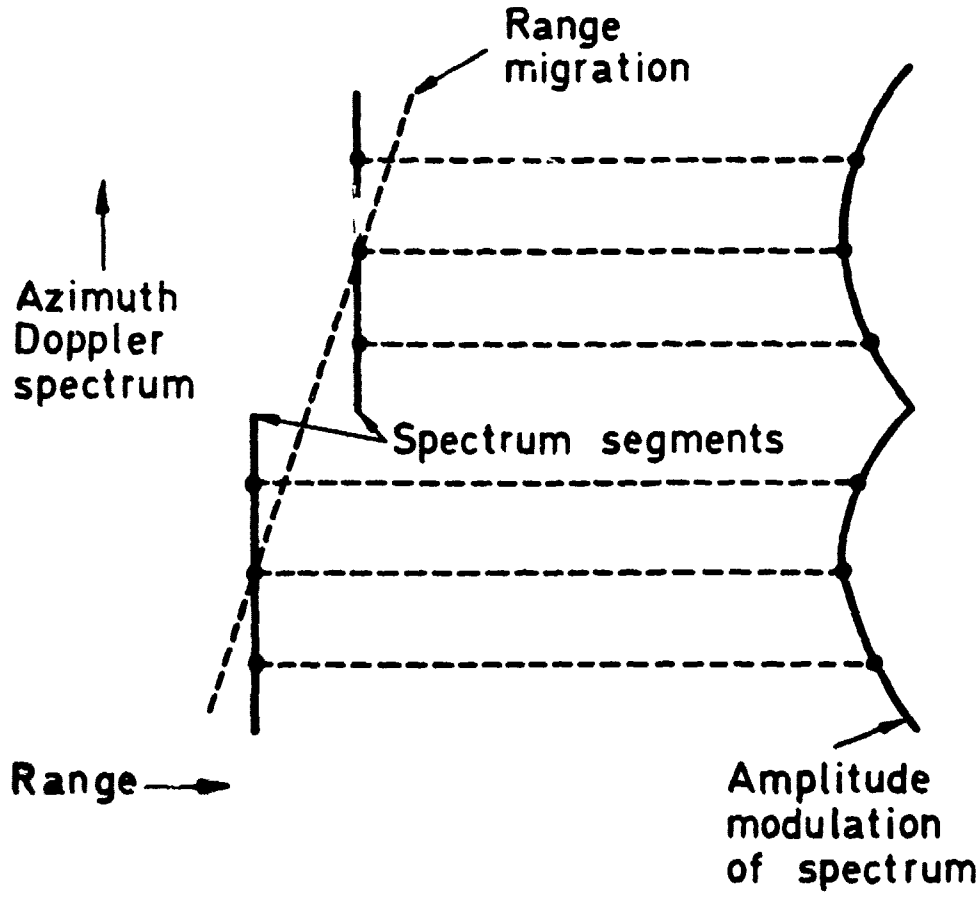

Fig 28 Amplitude modulation of the spectrum

REPORT DOCUMENTATION PAGE

Overall security classification of this page

UNCLASSIFIED

As far as possible this page should contain only unclassified information. If it is necessary to enter classified information, the box above must be marked to indicate the classification, e.g. Restricted, Confidential or Secret.

1. DRIC Reference (to be added by DRIC)	2. Originator's Reference	3. Agency Reference	4. Report Security Classification/Marking
	RAE TR 83079	N/A	UNCLASSIFIED

5. DRIC Code for Originator	6. Originator (Corporate Author) Name and Location
7673000W	Royal Aircraft Establishment, Farnborough, Hants, UK
5a. Sponsoring Agency's Code	6a. Sponsoring Agency (Contract Authority) Name and Location
N/A	N/A

7. Title

Theory of digital imaging from orbital synthetic aperture radar

7a. (For Translations) Title in Foreign Language

7b. (For Conference Papers) Title, Place and Date of Conference

8. Author 1. Surname, Initials	9a. Author 2	9b. Authors 3, 4	10. Date	Pages	Refs.
Barber, B.C.			November 1983	74	26

11. Contract Number	12. Period	13. Project	14. Other Reference Nos.
N/A	N/A		Space 631

15. Distribution statement

(a) Controlled by –

(b) Special limitations (if any) –

16. Descriptors (Keywords) (Descriptors marked * are selected from TEST)

Synthetic aperture radar. Synthetic antennas. Spaceborne SAR. Radar remote sensing. Digital imaging. Radar imaging.

17. Abstract

Digital synthetic aperture radar (SAR) imaging techniques have previously only been reported in the literature in a fragmentary manner. This article presents a comprehensive review of the theory of digital SAR imaging from Earth orbiting satellites. The digital SAR imaging process is explained, including a discussion of various aspects which are specific to satellite-borne SAR. A number of relevant digital processing techniques are reviewed and it is shown how these techniques may be applied to the processing of digital SAR data. The range migration problem is discussed and various techniques for overcoming it are presented. The paper should be useful not only to the designer of SAR processors, but also to the user of digital SAR data, and images.

910/1